MIHAI DINU

# PERSONALITATEA LIMBII ROMÂNE

ALL

Scanned with OKEN Scanner

MIHAI DINU

# PERSONALITATEA LIMBII ROMÂNE

## fizionomia vocabularului

ALL

Cartea Românească

1996

Scanned with OKEN Scanner

## 1. Vocabularul ca sistem: formularea problemei

S-a spus, nu o dată, despre vocabular că reprezintă compartimentul cel mai slab structurat al limbii. Comparativ, de pildă, cu familia relativ restrînsă, dar exemplar de închegată, a fonemelor, colecţia eteroclită de elemente de varii proveniențe ce constituie tezaurul lexical al unei limbi se prezintă, cel puţin la prima vedere, ca o imensă masă aluvionară sedimentată haotic, potrivit capriciilor acelor viituri istorice care au reuşit să forţeze porţile limbii, penetrînd-o în zona sa de (aparent) minimă rezistenţă.

Aflat în mai mare măsură „sub vremi" decît celelalte subsisteme ale limbii, vocabularul s-a dovedit într-adevăr dintotdeauna mult mai permeabil la influenţele alogene. Contactul său nemijlocit cu realităţile schimbătoare ale vieţii i-a impus un ritm de primenire de care atît fonologia cît şi morfologia se puteau dispensa. Menite, între altele, să asigure persistenţa specificului limbii, ferind-o de primejdia dizolvării în masa lingvistică ambiantă, acestea din urmă aveau nevoie tocmai de un anumit grad de etanşeitate şi inerţie care să le garanteze conservarea individualităţii într-o lume a prefacerilor necontenite. De aceea rareori de-a lungul istoriei ele au cedat presiunilor din afară, evoluţia lor avînd aproape întotdeauna motivaţii interne, iar atunci cînd, contrar tendinţei dominante, s-au văzut totuşi obligate să accepte



Scanned with OKEN Scanner

imixtiunea forţelor asediatoare, infiltrarea „inamicului" a avut loc (se putea altfel?) tot prin poarta mereu deschisă a vocabularului[1].

Dacă gramatica reprezintă un analog lingvistic al sistemului osos, schelărie interioară ce structurează şi susţine masa corporală a organismelor vii, asigurîndu-le stabilitatea indispensabilă supravieţuirii, vocabularul ne apare mai curînd ca un corespondent al plămînului, responsabil de oxigenarea tuturor celorlalte organe. Nimic nu ajunge să afecteze structurile lingvistice vitale înainte de a trece prin anticamera sistemului lexical, acolo unde se joacă în fond soarta viitoare a limbii. Confruntat direct cu schimbările din societate, vocabularul este cel dintîi care trebuie să se adapteze noilor situaţii, să răspundă creator, dar şi cu suficientă promptitudine, solicitărilor exterioare. De capacitatea sa de asimilare a noului depind în mod decisiv măsura evoluţiei sau a stagnării limbii, devenirea ori extincţia acesteia.

Mai dinamic, mai receptiv la stimuli externi, mai permeabil decît oricare alt sector al limbii, lexicul a putut părea insuficient structurat numai printr-o asociere arbitrară a ordinii cu rigiditatea, a organizării cu imobilismul. În realitate, experienţa altor domenii ale cunoaşterii demonstrează că adaptabilitatea, supleţea şi un metabolism mai intens sînt tocmai semne distinctive ale sistemelor superior organizate. Capacitatea vocabularului de a efectua transferuri operative din şi în „mediul ambiant", asimilînd uşor noutăţile, dar şi eliminînd rapid din fondul său activ balastul elementelor caduce ar sugera, în această perspectivă, că departe de a constitui o masă nediferenţiată, amorfă, lexicul se prezintă, dimpotrivă, ca un sistem complex, coerent şi fiabil totodată, elastic şi stabil.

Formulată aici ca o simplă conjectură, bazată pe analogii oricînd contestabile, ipoteza naturii superior organizate a sistemului lexical îşi află numeroase confirmări în achiziţiile recente ale lexicologiei, disciplină lingvistică străveche, aflată actualmente în plin proces de

---

[1] Este ceea ce s-a întîmplat, pentru a ne limita la un singur exemplu, cu readoptarea consoanei *h*, pătrunsă în româna comună prin mijlocirea aportului lexical slav şi consolidată graţie împrumuturilor ulterioare din turcă *(hain, huzur)*, maghiară *(ham, hotar)*, neogreacă *(heruvic, hrisov)* sau germană *(halbă, haltă)*. Cuvinte de origine slavă comune dialectelor româneşti nord- şi suddunărene ca *duh* sau *hrăni* atestă vechimea fenomenului. Să menţionăm, fără a-i acorda însă un credit exagerat, şi ipoteza lui Gr. Brâncuş *(Originea consoanei* h *din limba română* în SCL XII, 1961, pp. 471—475) potrivit căreia originea *h*-ului românesc trebuie căutată în influenţa substratului.



Scanned with OKEN Scanner

reîmprospătare, atît a repertoriului metodologic, cît şi a viziunii generale asupra propriului obiect de studiu. Ideea complexităţii structurale a vocabularului, multă vreme obnubilată de nivelul relativ scăzut de cunoaştere a intercondiţionărilor ascunse vederii dintre unităţile sale componente, se afirmă tot mai viguros, pe măsură ce studiilor referitoare la relaţiile „orizontale" dintre cuvinte, cum ar fi sinonimia sau antonimia, în care termenilor implicaţi li se atribuie o importanţă egală ori comparabilă în procesul comunicării, vin să li se alăture cercetări ce îşi propun stabilirea de ierarhii „pe verticală", bazate pe criterii ca polisemia, productivitatea lexicală sau frecvenţa de folosire. Prin încercarea ei de a extrage, din examinarea sub un unghi inedit a lexicului românesc, cîteva noi argumente în favoarea unei înţelegeri sistemice a vocabularului, lucrarea de faţă se subsumează şi ea direcţiei menţionate.

## 2. Frecvenţa cuvintelor şi problemele determinării ei

Că, sub raportul uzajului, nu toate cuvintele sînt la fel de importante o afirma, încă de acum un veac, teoria hasdeiană a „circulaţiei cuvintelor". Metoda sondajelor de frecvenţă pe care, cel dintîi la noi, a folosit-o marele lingvist, pentru a dovedi netemeinicia afirmaţiilor lui A. de Cihac[1] privind repartiţia pe origini a vocabularului românesc, a făcut de atunci o frumoasă carieră. Nu puţini cercetători au considerat-o chiar drept modalitatea cea mai

[1] Îi revine lui Mircea Seche meritul de a fi reconsiderat cu obiectivitate şi competenţă pe nedrept hulitul *Dicţionar etimologic român* al lui A. de Cihac, lucrare ce „inaugurează un curent de cercetare ştiinţifică modernă, în general străin de ideile preconcepute" (*Schiţă de istorie a lexicografiei române* Editura Ştiinţifică, 1966), ridicînd, după aproape un secol, anatema aruncată prin execuţia nedreaptă la care a supus-o Hasdeu. Numărătoarea efectuată de M. Seche a relevat, de altfel, o surprinzătoare neconcordanţă între procentele repartiţiei pe origini a materialului lexical din *Dicţionar* ce au provocat reacţia violentă a lui Hasdeu (40% cuvinte slave, faţă de numai 20% latine, cifre lansate, fără o verificare riguroasă, de însuşi Cihac (!)) şi situaţia reală, arătînd că, de fapt, 45% dintre cuvinte sînt de origine latină şi numai 35% de provenienţă slavă.



Scanned with OKEN Scanner

obiectivă de evaluare a importanţei cuvintelor. Este suficient să amintim, bunăoară, că atunci cînd, inspirat de un binecunoscut model străin[1] al anilor '50, Al. Graur a elaborat primul proiect de vocabular fundamental al limbii române[2] (fondul principal lexical), criticile de principiu ce s-au adus demersului său se întemeiau tocmai pe ideea (prejudecata?) că ordinea reală de importanţă a cuvintelor nu ne-o poate furniza decît statistica. În recenzia pe care o făcea noii versiuni[3], din 1957, a lucrării lui Al. Graur, I. Fischer[4] afirma că: „Prin statistica frecvenţei se poate ajunge la o ierarhizare a cuvintelor din limbă, problema stabilirii limitelor fondului principal rămînînd însă deschisă. Dacă o asemenea limită există într-adevăr în mod obiectiv, ea ar putea apărea în statistică sub forma unei bruşte scăderi a frecvenţei, pornind de la un anumit număr de cuvinte. Dacă însă trecerile sînt lente, şi înclinăm să credem că aşa stau lucrurile, urmează să examinăm dacă concluziile care se pot trage (şi care au fost trase de acad. Al. Graur) din studiul fondului principal sînt sau nu caduce".

Nu este locul aici şi nici nu ne propunem să intrăm în detaliile unei dispute astăzi clasate, care a mai antrenat şi alte nume ale lingvisticii româneşti, luînd la un moment dat o turnură destul de delicată (oricît ar părea de ciudat astăzi, din adeziunea la ideea de fond principal lexical se făcea în epocă un criteriu de verificare a fidelităţii ideologice a cercetătorului...), dar trebuie spus că stabilirea inventarelor de frecvenţă nu s-a dovedit nici pe departe panaceul pe care îl întrezărea atunci I. Fischer. După o perioadă de entuziasm ce nu a depăşit finele deceniului al şaptelea, interesul pentru statistica lexicală a înregistrat o scădere simţitoare de îndată ce au devenit perceptibile dificultăţile de extrapolare a rezultatelor obţinute pe un corpus limitat de texte la nivelul întregii limbi sau măcar al uneia dintre variantele ei stilistice. Pe cît de semnificativă este statistica pentru profilul lexical al materialului excerptat, pe atît de înşelătoare se dovedeşte ea cînd i se cere să dea seamă de trăsăturile generale ale ansamblului limbii. O paranteză este aici necesară.

Cu excepţia unui număr neglijabil de situaţii atipice, rostul limbajului este acela de a transmite un anumit conţinut. Vorbim sau

---

[1] I.V. Stalin: *Marxismul si problemele lingvisticii,* Editura pentru literatură politică, 1953.

[2] *Încercare asupra fondului principal lexical al limbii române,* Editura Academiei R.P.R., 1954.

[3] *Fondul principal al limbii române,* Editura Academiei R.P.R., 1957.

[4] *Al. Graur: Fondul principal al limbii române (recenzie),* SCL IX, 1958, pp. 127—133.



Scanned with OKEN Scanner

scriem pentru a comunica ceva, iar alegerea cuvintelor pe care le folosim depinde în mod esenţial de natura mesajului vehiculat. Sfera semantică a comunicării dictează selecţia lexicală în aşa măsură încît, dacă facem abstracţie de cuvintele instrumente gramaticale (prepoziţii, conjuncţii, pronume), a căror utilizare este inevitabilă, găsirea a două texte total disjuncte sub raportul vocabularului folosit nu prezintă nici o dificultate. Instrucţiunile de folosire a unui medicament şi cele ale unei maşini de spălat rufe posedă rareori cuvinte „pline" în comun. Nu altfel stau lucrurile şi în domeniul comunicării orale, unde cu alte cuvinte comentăm un meci de fotbal şi cu altele o expoziţie de pictură.

Această specializare a lexicului explică de ce nu este necesară nici o statistică, ba nici măcar deschiderea cărţilor respective, pentru a anticipa că într-un tratat de oftalmologie frecvenţa cuvîntului „ochi" va fi cu mult mai mare decît într-un manual de tehnologie a construcţiilor de maşini şi că, în mod aproape cert, vocabule ca *dor* sau *crîng* nu vor putea fi găsite nici într-unul, nici în celălalt. Dacă ultimele două cuvinte sînt larg răspîndite în poezie, ar fi în schimb inutilă despuierea întregii lirici româneşti în căutarea cuvîntului *pancreas,* a cărui prezenţă acolo este, o ştim bine, cu totul improbabilă.

Profilul lexical atît de deosebit al unor texte reprezentînd stiluri ale limbii şi domenii de preocupare diferite ridică o problemă care constituie piatra de încercare (dar şi de poticnire) a statisticii lingvistice: cum trebuie procedat pentru a determina frecvenţa cuvintelor nu într-un text sau altul, ci în limbă în general? Numărarea directă a ocurenţelor fiecărui cuvînt din tot ce s-a scris şi vorbit vreodată în româneşte fiind irealizabilă, sîntem constrînşi să recurgem la alegerea unui număr de texte care să alcătuiască, pe cît posibil, o selecţie reprezentativă. Nu este însă greu de văzut că, printr-o atare alegere, se pierde tocmai elementul ce constituie punctul forte al statisticii şi anume acela de a fi *obiectivă.* Raporturile numerice dintre cuvintele *ochi* şi *strung* vor depinde, între altele, şi de lungimea textelor din domeniul oftalmologiei şi, respectiv, al construcţiilor de maşini cuprinse în eşantionul considerat. Opţiunea celui ce a operat selecţia textelor devine astfel un factor determinant în stabilirea scării de frecvenţă a cuvintelor. Accentul se deplasează de la obiectivitatea inventarului impersonal către arbitrariul unei decizii esenţialmente subiective.

Supuşi acestei fatalităţi, lingviştii care s-au încumetat să încerce o estimare a frecvenţei cuvintelor româneşti s-au văzut nevoiţi să rezolve fiecare în felul său problema ponderii în corpusul utilizat a eşantioanelor de limbă aparţinînd diferitelor domenii specializate. Astfel,



Scanned with OKEN Scanner

dicţionarul de frecvenţe al lui Alphonse Juilland[1] a fost întocmit pe baza despuierii unui corpus de 500 000 de cuvinte-text, reprezentînd în mod egal (cîte 100 000 pentru fiecare): teatrul, proza beletristică, memorialistica, presa şi literatura tehnică. Să observăm că, deşi poezia lipseşte, ponderea acordată stilului artistic o depăşeşte cu 50% pe cea a stilului zis „ştiinţific-administrativ" (în care se include şi publicistica). În schimb, statistica datorată lui Constant Maneca[2], în ciuda faptului că acordă şi poeziei locul cuvenit, modifică proporţia în favoarea celui de al doilea stil al limbii, care, cu 24 000 de cuvinte-text, se apropie ca întindere de spaţiul atribuit literaturii beletristice (26 000 de cuvinte-text). De partea cui este dreptatea, dacă de dreptate poate fi vorba? Oricum am lua-o, chestiunea rămîne indecidabilă.

Pe de altă parte, nici limitarea cercetării la un singur stil al limbii literare nu simplifică prea mult problema. Gh. Bolocan a consacrat o asemenea cercetare[3] stilului publicistic, iar un colectiv de lingvişti grupaţi în Cercul de lingvistică matematică de pe lîngă Secţia de fonetică şi dialectologie a Institutului de lingvistică din Bucureşti a efectuat, încă din anul 1958, la iniţiativa lui Al. Rosetti şi sub conducerea lui Valeriu Şuteu, o statistică[4] a lexicului literaturii beletristice, bazată pe excerptarea unui material totalizînd 50 000 de cuvinte-text din poezia (T. Arghezi, M. Codreanu, O. Goga), proza (M. Sadoveanu, I. Agârbiceanu, I. Al. Brătescu Voineşti, E. Camilar, L. Rebreanu, M. Preda) şi teatrul (Camil Petrescu) secolului nostru. Discuţii nesfîrşite şi, evident, oţioase se pot purta în legătură cu gradul de „reprezentativitate" al acestei din urmă selecţii. Dar dacă sîntem poate îndreptăţiţi să ne întrebăm de ce din vasta operă sadoveniană s-a preferat romanul *Nicoară Potcoavă* sau de ce bogata dramaturgie a epocii este prezentă numai prin piesele lui Camil Petrescu, ne revine atunci şi obligaţia de a oferi o soluţie alternativă, iar aceasta va suferi inevitabil de acelaşi viciu ca şi opţiunea incriminată. Alegerea noastră, prin însuşi faptul că e a noastră, nu va fi cu nimic mai obiectivă decît cea a lui Valeriu Şuteu sau a oricui altcuiva.

---

[1] Alphonse Juilland, P.M.G. Edwards, Ileana Juilland: *Frequency dictionary of romanian words,* Mouton et comp., London, The Hague, Paris, 1965.

[2] C. Maneca: *Consideraţii asupra frecvenţei cuvintelor în limba română literară contemporană,* SCL XVII nr. 6, pp. 623—633, Bucureşti 1966.

[3] Gh. Bolocan: *Unele caracteristici ale stilului publicistic al limbii române literare,* SCL XII, nr. 1, pp. 35—73, Bucureşti, 1961.

[4] V. Şuteu: *Observaţii asupra frecvenţei cuvintelor în operele unor scriitori români,* SCL X, nr. 3, pp. 419—445, Bucureşti, 1959.



Scanned with OKEN Scanner

În plus, amestecarea într-un inventar comun a preferinţelor lexicale ale unor scriitori atît de personali ca Arghezi, Sadoveanu sau Marin Preda prezintă şi un alt inconvenient major: ea nu numai că oferă o imagine mult prea aproximativă, trunchiată şi deci înşelătoare asupra a ceea ce este limba literaturii române în general, dar şi anulează prin diluare într-o pastă nediferenţiată particularităţile de limbă ale fiecărui creator în parte. Soluţie funciarmente hibridă, statistica astfel întocmită rămîne cantonată în zona incertă a jumătăţilor de măsură, ea are un caracter prea particular pentru a da seamă de situaţia de ansamblu a limbii şi, totodată, prea general pentru a fi de ajutor cercetărilor de stilistică. Acestea din urmă nu pot profita realmente decît de mai restrînsele, dar mai semnificativele analize consacrate scrisului unui singur autor.

Deosebit de relevante, insuficient valorificate însă, cel puţin deocamdată, de exegeza literară, s-au dovedit în acest domeniu studiile consacrate vocabularului eminescian. Pentru scepticii tentaţi să minimalizeze importanţa metodelor cantitative e suficient să amintim, de exemplu, că, potrivit statisticii întocmite de Luiza Seche[1], cele mai frecvente 4 substantive din poezia şi proza artistică eminesciană sînt, în ordine: *ochi, lume, viaţă, umbră*. Ne îndoim că ermeneutica literară a reuşit să ofere pînă în prezent o caracterizare mai adîncă şi, totodată, mai lapidară a universului de gîndire şi a climatului liric al poeziei lui Eminescu decît cea pe care o cuprind aceste numai patru cuvinte:

„Ochi — lume, viaţă, ... umbră!"

Un poem-sinteză care, într-un singur vers, exprimă esenţialul despre atitudinea de contemplator dezamăgit al realităţii pe care ne-o relevă opera poetului! De n-ar fi decît acest rezultat emoţionant, şi tot am putea socoti că munca de benedictin a autoarei statisticii nu s-a irosit în zadar...

Din cele spuse ar putea să se desprindă însă concluzia, nedorită de noi, căci sîntem departe de a gîndi astfel, că relevanţa statisticii ar fi condiţionată de exhaustivitatea ei. Pentru a fi cu adevărat semnificativă i s-ar cere aşadar să îmbrăţişeze în totalitate obiectul studiat. Că lucrurile nu stau aşa ne-o dovedeşte cu prisosinţă practica îndelungată a atîtor domenii, printre care şi lingvistica însăşi, care au avut de profitat şi continuă să profite de pe urma sondajelor statistice.

---

[1] Luiza Seche: *Lexicul artistic eminescian în lumină statistică*, Editura Academiei R.S.R., Bucureşti, 1974.



Scanned with OKEN Scanner

Utilitatea abordării unei mulţimi numeroase prin intermediul numai al unei submulţimi relativ restrînse a ei este astăzi în afara oricărei discuţii. Ceea ce însă poate fi şi trebuie neapărat discutat sînt limitele de aplicare ale unei atari metode. Cu alte cuvinte, se impune să precizăm care tipuri de probleme sînt susceptibile de a-şi găsi rezolvarea prin extrapolarea unor rezultate obţinute pe eşantioane limitate şi care nu.

Analogia ne este şi aici de folos. Măsurarea staturii a 1 000 de bucureşteni judicios aleşi ne poate furniza, cu o aproximaţie acceptabilă, înălţimea medie a locuitorilor Capitalei. Ea nu ne spune însă nici cît măsoară cel mai înalt bucureştean, rămas probabil în afara eşantionului considerat, şi nici, cu atît mai puţin, cine este acesta. Tot astfel, o mostră de limbă scrisă sau vorbită, de întindere îndestulătoare, ne poate îngădui să estimăm frecvenţa relativă în româneşte a substantivelor sau verbelor, dar ea rămîne mută în privinţa locului pe care cutare substantiv sau verb îl ocupă în lista celor mai folosite cuvinte româneşti. Prin analiza unui eşantion corect selecţionat se pot deci preciza numai anumiţi parametri statistici globali ai colectivităţilor analizate, dar nu şi coordonatele concrete ale unui membru anume al acestora. Ordonarea pe criteriul uzajului a întregului vocabular, în măsura în care presupune cunoaşterea precisă a poziţiei în sistem a unor cuvinte individuale, rămîne aşadar în afara posibilităţilor metodei sondajelor. Ea ar necesita o statistică completă a limbii scrise şi vorbite, de azi şi dintotdeauna, deziderat a cărui transpunere în fapt nici măcar nu o întrezărim.

## 3. Locul frecvenţei de folosire printre ceilalţi indicatori ai importanţei cuvintelor

Odată ajunşi la concluzia că ordonarea completă a vocabularului potrivit frecvenţei de folosire a cuvintelor este cvasiimposibilă, ne putem întreba, ca vulpea din zicală, dacă nu cumva, poamele jinduite erau de fapt acre. Altfel spus, s-ar putea ca studiul limbii să nici nu aibă nevoie de un asemenea clasament complet al unităţilor lexicale. Efortul depus pentru întocmirea lui ar fi, în acest caz, inutil.



Scanned with OKEN Scanner

Să observăm mai întîi că frecvenţa de folosire a cuvintelor depinde în mod direct de caracterul mai mult sau mai puţin uzual al noţiunilor pe care acestea le desemnează. Punînd semnul egalităţii între cuvinte şi lucruri (asimilare forţată, dar nu întrutotul arbitrară, care de altfel s-a mai făcut, ba chiar a fost ridicată de o anumită şcoală lingvistică la rangul de principiu metodologic), ar fi de aşteptat ca, practic în toate limbile pămîntului, cuvintele cele mai frecvente să fie cele care denumesc aceleaşi aspecte ale realităţii şi anume cele indispensabile vieţii oamenilor de pretutindeni.

Chiar dacă lucrurile nu stau întocmai aşa, este evident că priorităţile statistice ale limbii reflectă în mare măsură priorităţile existenţei umane, anumite cuvinte revenind mereu în comunicarea lingvistică fiindcă se referă la obiecte sau situaţii de care ne lovim mai des, iar nu fiindcă ar reprezenta termeni mai importanţi ai sistemului lexical. Dar o noţiune fundamentală nu este redată în mod necesar printr-un cuvînt fundamental, căci, deşi strîns legate între ele, limba şi lumea pe care aceasta o transpune în cuvinte constituie realităţi guvernate de legi deosebite.

Vîntul şi ploaia sînt fenomene între care, în ordinea relaţiei om-natură nu se poate stabili nici un raport de întîietate, însă într-o limbă sau alta, cuvintele ce le desemnează pot ocupa, fie unul, fie celălalt, un loc privilegiat. Astfel în româneşte cuvîntul *vînt* domină cuvîntul *ploaie,* atît din punctul de vedere al productivităţii lexicale (în DLRM sînt consemnate 23 de derivate ale sale, faţă de numai 6 ale lui *ploaie*), cît şi din cel al numărului de sensuri distincte (în DEX: 4 faţă de 2) şi al varietăţii expresiilor şi locuţiunilor la care participă (11 faţă de 3 în DEL). Acest raport este propriu limbii române şi nu există nici un temei de a afirma că lucrurile trebuie să stea la fel şi în alte idiomuri. În schimb, vorbitorii din multe alte regiuni ale lumii (cel puţin din cele cu climat asemănător cu al nostru) au prilejul să pomenească în conversaţiile lor cele două fenomene meteorologice cam la fel de des ca şi românii. Faptul că frecvenţa de folosire depinde de o realitate situată în afara limbii şi pe care limba nu o poate influenţa face ca acest parametru să prezinte o importanţă lingvistică mai redusă. El are defectul de a amalgama criteriile, nelăsînd să se distingă limpede aportul planului lingvistic în situarea cuvîntului pe o anumită poziţie în preferinţele vorbitorilor. E greu de spus cît din frecvenţa ridicată a cuvîntului *pîine* se datoreşte însemnătăţii vitale a obiectului pe care îl desemnează şi



Scanned with OKEN Scanner

cît locului ocupat de acest cuvînt în angrenajul sistemului lexical[1] românesc.

În general, operarea distincţiei dintre planul lingvistic şi planul realităţii extralingvistice este mai puţin lesnicioasă decît pare. La întrebarea ce categorii de cuvinte aparţin vocabularului fundamental, lingviştii răspund adesea enumerînd realităţi curente ale existenţei, în genul: „cuvinte care denumesc părţile corpului, îmbrăcămintea, locuinţa, hrana şi modul ei de preparare etc.". Însuşi modul de formulare a unui asemenea răspuns dezvăluie confuzia planurilor. Este evident că, deşi vorbeşte despre *cuvinte* fundamentale, cel întrebat are de fapt în vedere *noţiunile* fundamentale, ceea ce este cu totul altceva. Stabilirea acestora din urmă nu priveşte lingvistica, în ciuda faptului că între noţiuni şi cuvinte există o strînsă legătură. În orice caz o noţiune fundamentală nu este exprimată în mod necesar printr-un cuvînt fundamental. Existenţa sinonimiei are un rol însemnat în apariţia acestei discrepanţe. Faptul că mai multe unităţi lexicale diferite sînt nevoite să îşi împartă sarcina desemnării unei aceleiaşi noţiuni le slăbeşte implicit poziţia în sistemul vocabularului. Importanţa lor în limbă apare atunci mult diminuată în raport cu importanţa noţiunii exprimate. În zadar denumeşte cuvîntul *brîncă* o parte esenţială a trupului omenesc: el rămîne un termen cu totul periferic al vocabularului, tocmai din cauza poziţiei incomparabil mai puternice a unor sinonime ca *mînă* sau *braţ*. Nici situaţia inversă, în care unei noţiuni puţin însemnate i se asociază un cuvînt bine situat în ierarhiile limbii, nu este deloc rară. Un exemplu, citat de Al. Graur[2], este cel al cuvîntului *doagă*, termen vechi şi productiv (a dat derivate ca *dogar, dogărie, dogi, dogit*), prezent totodată în cîteva expresii binecunoscute, dar care denumeşte un obiect deloc indispensabil traiului.

Pentru punerea în evidenţă a autonomiei reflectării lingvistice în raport cu realitatea reflectată este metodologic utilă aplicarea procedeului, curent în alte ştiinţe, al „separării efectelor". Analizînd bunăoară poziţia relativă a unor cuvinte ce redau aspecte ale existenţei de importanţă sensibil echivalentă, se elimină efectul perturbator pe care deosebirea de însemnătate dintre noţiuni îl produce asupra ponderii cuvintelor în limbă. Să considerăm un exemplu simplu.

---

[1] Cuvîntul *pîine* numără puţine derivate (exclusiv diminutive: *pîinică, pîinişoară, pîiniţă*), dar are sensuri figurate (ca cel de *slujbă* sau *funcţie*) ori extinse metonimic, prin procedeul *pars pro toto* („pîinea cea de toate zilele") şi intră în alcătuirea unor expresii mult folosite *(a avea pîinea si cuţitul, a se vinde ca pîinea caldă, a lua pîinea de la gură, a fi bun ca pîinea caldă* etc.).

[2] *Încercare* ... pag. 20



Scanned with OKEN Scanner

La latitudinea României cele patru anotimpuri sînt practic de durată egală şi la fel de importante pentru viaţa locuitorilor. Nu ar exista deci motive de ordin obiectiv ca vreunul dintre ele să se bucure de un regim preferenţial în reflectarea lingvistică. Cu toate acestea, o simplă consultare a dicţionarului relevă deosebiri frapante de poziţie în sistemul lexical a cuvintelor *iarnă, primăvară, vară* şi *toamnă*. Primul dintre ele, deşi susţinut de *iernă* şi *iernat,* care descind însă dintr-un etimon latin diferit, nu are în DLRM decît 3 derivate (subst. *iernatic* = adăpost de iarnă, adj. *iernatic* = caracteristic iernii şi subst. *iernător* = bazin de iernat pentru peşti), faţă de 9 ale lui *toamnă (tomnă, tomnat, tomnatic 1, tomnatic 2, tomniţă, tomniu, întomna, întomnare, întomnat).* Deşi formarea sa este de dată latină, *primăvară,* cuvînt ce posedă la rîndul său o familie numeroasă *(primăvărat, primăvăratic, primăvăriţă, primăvăros, desprimăvăra, desprimăvărare, desprimăvărat, împrimăvăra),* e încă simţit drept un derivat al lui *vară,* care a mai generat pe *văra, vărat 1, vărat 2, văratic 1, văratic 2* şi *văruţă.* Aşadar reflectarea în limba română a celor patru anotimpuri are loc inegal, sub forma a trei cuiburi lexicale de extensii sensibil diferite: cuibul ⟨*vară, primăvară*⟩ care conţine 16 termeni, cuibul ⟨*toamnă*⟩, alcătuit din 10 elemente şi cuibul ⟨*iarnă, ierna*⟩ ce cuprinde numai 6 cuvinte. Inexistenţa sinonimelor face ca, în acest caz, deosebirile să nu poată fi puse pe seama concurenţei acestora. Ele au probabil alte cauze, pe care însă nu ne propunem să le investigăm aici, cu atît mai mult cu cît ne-ar purta departe de subiectul nostru, pe tărîmul unei explorări lingvistice a Weltanschauung-ului românesc de felul celei practicate în trecut, cu ingeniozitate şi talent, de C. Noica.

Studiul comparativ al unor antonime ne oferă de asemenea prilejul de a evidenţia gradele de libertate ale organizării sistemului lexical faţă de structurile realului. Termenii unor opoziţii ca *mare* vs. *mic, greu* vs. *uşor, lung* vs. *scurt,* fiind definiţi în mod relativ, prin comparaţie, nici unul dintre ei nu este favorizat a priori, în virtutea vreunei superiorităţi de ordin obiectiv. Lucrurile, fiinţele, intervalele de timp sau distanţele mari nu sînt mai numeroase ori mai puţin numeroase decît cele mici, deoarece mărimea nu constituie o calitate în sine, raportabil la un etalon absolut, şi, ca atare, un acelaşi lucru poate fi considerat (şi deci numit) alternativ *mare* sau *mic,* după împrejurări. Există, prin urmare, un anumit echilibru conceptual între cei doi poli ai relaţiei antonimice, dar el nu se traduce în mod obligatoriu printr-un raport de forţe echilibrat între cuvintele respective. Adesea, unul dintre termeni a reuşit să pătrundă mai adînc în limbă, să-şi creeze o reţea mai întinsă de legături, o familie mai



Scanned with OKEN Scanner

puternică. Este cazul adjectivului *lat* care, deşi la fel de vechi ca şi antonimul său *îngust* (aparţin ambele fondului latin moştenit), a generat un număr mult superior de derivate[1] *(lateş, lătăreţ, lăţi, lăţime, lăţire, lăţiş, lăţit, răzleţi, răzleţire, răzleţ)* şi participă la expresii idiomatice ca: *a cădea lat, a avea limbă lată, a lăsa (pe cineva) lat, a spune vorbe mari şi late* (DEL pp. 130—131). Alteori antonimul aparţine chiar familiei lexicale a corelativului său[2]: *nevinovat* provine din *vinovat* contribuind, alături de alte derivate ca *vinovăţie, învinovăţi, învinovăţire, nevinovăţie, dezvinovăţi, dezvinovăţire* la consolidarea poziţiei în sistem a termenului matcă. O situaţie limită e cea în care unul dintre polii opoziţiei nici nu dispune de o expresie lingvistică proprie, cum se întîmplă cu antonimul lui *adînc,* cuvînt care există în limbi ca sîrbocroata, rusa ori slovaca[3], dar nu şi în româneşte, unde se exprimă perifrastic sub forma *puţin adînc*. În acest caz absenţa unui lexem ad-hoc se repercutează asupra frecvenţei termenului pozitiv, care îşi dublează practic ocurenţele, obligat să participe şi la exprimarea noţiunii opuse.

Se vede din exemplele de mai sus cît de complexă este relaţia dintre frecvenţă şi ceilalţi parametri definitorii pentru importanţa cuvintelor în limbă. O simplă carenţă de productivitate precum cea a cuvîntului *adînc,* incapabil să genereze un antonim de tipul *neadînc,* obligă vorbitorii la folosirea mai frecventă a celui dintîi. Polisemia sporeşte şi ea frecvenţa, întrucît a utiliza un termen în mai multe accepţiuni înseamnă a-l utiliza mai des. Cît despre calitatea de a intra în expresii şi locuţiuni, este de la sine înţeles că se reflectă şi ea asupra frecvenţei, în măsura în care formulele respective se bucură de circulaţie.

Dependenţa parametrului frecvenţă de ceilalţi indicatori lexicali, principial incontestabilă, este departe de a fi cunoscută în detaliu, dar, dincolo de aspectele încă insuficient elucidate ale problemei, trebuie spus că stabilitatea cuvintelor în limbă o datorăm mai puţin celei dintîi decît celor din urmă. Un cuvînt este cu atît mai greu de dizlocat din

---

[1] Să observăm, şi asupra acestui aspect vom mai reveni, că productivitatea lexicală constituie un indicator al vechimii numai în medie, pentru ansamblul cuvintelor de o anumită origine. La nivel individual criteriul nu mai operează. Aşa se face că autohtonul *ceafă* nu are descendenţi, pe cînd neologismul *alimenta* posedă un număr apreciabil de derivate: *alimentară, alimentare, realimenta, subalimenta, subalimentare, subalimentat, supraalimenta, supraalimentare, supraalimentat.*

[2] Se vorbeşte în acest caz de *antonime homolexe* (v. R. Sîrbu: *Antonimia lexicală în limba română,* Editura Facla, Timişoara, 1977).

[3] R. Sîrbu, *op. cit.,* pag. 112.



Scanned with OKEN Scanner

sistemul lexical cu cît şi-a creat un grup mai puternic de „susţinători", constînd în înţelesuri, expresii şi derivate diferite, toate prilejuri de a-l evoca necontenit, ceea ce îl pune la adăpost de eventuala concurenţă a unui nou venit. În schimb, faptul de a nu fi decît des folosit constituie o garanţie insuficientă pentru menţinerea unui cuvînt în limbă, cum o probează nenumărate exemple de termeni fundamentali, înlocuiţi de-a lungul timpului cu împrumuturi mai recente.

Denumirile lunilor fac parte, fără dubiu, dintre cuvintele des şi unanim întrebuinţate. Rareori trece o zi fără să rostim numele vreunei luni, dar, cu toată această prezenţă continuă în conştiinţa vorbitorilor, termenii respectivi au constituit, probabil întotdeauna, elemente slabe, periferice, ale vocabularului, care, din această cauză, a acceptat uşor substituirea lor cu termeni noi. Dintre denumirile populare ale lunilor, ce au capitulat fără a opune prea multă rezistenţă în faţa terminologiei culte, nespecialistul nu le mai recunoaşte decît pe cele aparţinînd unor familii lexicale încă vii *(florar, cireşar, cuptor, brumar),* ignorîndu-le de regulă pe cele ce reprezentau elemente izolate, neproductive *(răpciune, undrea).* La rîndul lor, noile denumiri nu s-au dovedit nici ele apte de a-şi crea „ataşe" lexicale care să le garanteze perenitatea, astfel încît, dacă s-ar intenţiona aceasta, înlocuirea lor nu ar fi deloc imposibilă (romanii o făceau prin decrete, ca cele ce au rebotezat *(mensis) Iulius* sau *(mensis) Augustus,* lunile natale ale lui Caesar şi Octavianus).

Cu totul altfel se petrec lucrurile atunci cînd se încearcă substituirea cu un împrumut recent a unui cuvînt solid ancorat în profunzimile sistemului lexical. Un caz tipic ar fi, în această privinţă, concurenţa efemeră pe care neologismul *amor* a făcut-o, în secolul trecut, vechiului împrumut slav *dragoste.* Pe lîngă abundenţa derivatelor *(drăgăstos, drăgosti, drăgoste(a)lă, drăgostire, îndrăgosti, îndrăgostire, îndrăgostit)* şi a locuţiunilor *(a avea dragoste să, a se da în dragoste, a face dragoste, a prinde dragoste de, a se topi de dragoste)* acesta din urmă beneficia şi de sprijinul vecinătăţii semantice şi formale a rudei sale *drag,* posesoare, la rîndul ei, a unei bogate progenituri *(drăgălaş, drăgălăşenie, drăguţ, drăguleţ, drăguliţă, drăguş, îndrăgi, îndrăgit, îndrăgitor* etc). În aceste condiţii, era limpede că, în pofida eforturilor unora dintre cei mai însemnaţi creatori de limbă literară, dintre care nu a lipsit nici Eminescu, de a acredita noul termen, rivalitatea nu putea fi tranşată decît în favoarea cuvîntului mai vechi. Deşi legat pentru totdeauna de unele reuşite incontestabile ale poeziei româneşti („Ce e amorul?", „S-a dus amorul" etc.), nou-venitul nu a izbutit să depăşească pragul unui anumit, jargon,



Scanned with OKEN Scanner

poposind în cele din urmă undeva la periferia vocabularului, copleşit de conotaţia ironică pe care simţul lingvistic fără greş al lui Caragiale o intuia încă din vremea cînd ceilalţi clasici ai noştri îl tratau cu toată seriozitatea.

Un proces de sens contrar a condus în aceeaşi perioadă la marginalizarea slavului *norod,* element neproductiv, disparat, răzleţ, şi, ca atare, lesne de înlăturat, ce nu putea spera să reziste cu succes ofensivei latinului[1] *popor*. Rezultatul a fost statutul de arhaism la care a ajuns cuvîntul *norod* în raport cu limba română contemporană. Şansele ca un accident similar să i se întîmple reprezentantului unei familii bine instalate în limbă sînt, în schimb, cu mult mai reduse. Folosirea mai frecventă a neologismului *distruge,* care acoperă toate sensurile străvechiului *nimici* şi mai cunoaşte şi altele în plus *(mi s-a distrus încălţămintea, abuzurile i-au distrus sănătatea),* nu a fost suficientă pentru a-l elimina, căci *nimici* era vlăstarul puternicei familii a lui *nimic* şi poseda la rîndul său derivate *(nimicie, nimicire, nimicit, nimicitor)*.

Am conchide că, în raport cu caracteristicile curat lingvistice ca productivitatea lexicală, polisemia şi participarea la expresii, frecvenţa de întrebuinţare constituie un parametru mai puţin reprezentativ pentru locul ocupat de cuvînt în sistemul vocabularului. Dacă Gh. Bolocan, explorînd presa a trei decenii (1920—1939 şi 1950—1960) a găsit să substantivul cel mai des folosit este *partid,* acest rezultat ne spune, în fond, destul de puţin despre poziţia în limbă a cuvîntului *partid,* deşi este profund semnificativă pentru impactul pe care l-a avut realitatea numită *partid* în viaţa politică şi socială a ţării din perioada considerată. Pricini exterioare (din rîndul cărora nu este exclusă nici o anumită modă lingvistică) pot face ca frecvenţa unui cuvînt să crească sau să scadă, dar rolul său în limbă depinde relativ puţin de aceste fluctuaţii. Desigur, dispariţia unui cuvînt este precedată de o scădere a frecvenţei sale de utilizare, însă nu această scădere constituie cauza eliminării cuvîntului din limbă. Ea este doar un simptom al unei maladii a cuvîntului ce are întotdeauna un substrat lingvistic (redu-

---

[1] Fie că este moştenit, cum tind să creadă, în ciuda unor dificultăţi de ordin fonetic, I. Coteanu şi M. Sala (v. *Etimologia si limba română,* Editura Academiei R.S.R., 1987, pp. 90—93), fie că a fost reintrodus mai tîrziu pe cale cărturărească, cuvîntul *popor* are în limba română o vechime de cel puţin patru secole, posedînd atestări scrise încă din veacul al XVI-lea şi, datorită nobilei sale origini, s-a bucurat de tot sprijinul intelectualilor patrioţi din epoca renaşterii noastre naţionale.



Scanned with OKEN Scanner

cerea excesivă a corpului fonetic, o omonimie stînjenitoare, sterilitatea lexicală, combinată cu concurenţa unui sinonim mai puternic etc).

Înclinăm, de aceea, să dăm dreptate acelor lingvişti care, în căutarea elementelor pivot ale sistemului lexical, nu au apelat la criteriul frecvenţei decît ca mijloc auxiliar, colateral, de confirmare a importanţei unor cuvinte. Aşa a procedat Al. Graur în lucrarea amintită, realizînd o selecţie în genere judicioasă a celor mai reprezentative unităţi lexicale româneşti (fondul principal lexical). Expunerea considerentelor ce au stat la baza întocmirii listei de cuvinte atribuite fondului principal lexical (FPL) abia dacă include o referire la frecvenţă, căreia i se recunoaşte doar un oarecare rol, mai curînd negativ, în identificarea acelor unităţi ce *nu* aparţin FPL: „... cuvintele care se repetă foarte rar nu pot avea pretenţia să figureze în fondul principal lexical" (op.cit. pag. 39).

Criteriile de bază cu care a operat Al. Graur în stabilirea scării de importanţă a cuvintelor au fost: bogăţia de înţelesuri (polisemia), numărul de derivate (productivitatea) şi abundenţa expresiilor în care figurează. Ar mai fi de adăugat (de fapt în enumerarea autorului acest criteriu este cel dintîi) utilizarea generală, pe tot teritoriul ţării şi în toate straturile societăţii. Menit a elimina din FPL regionalismele, cuvintele de argou şi termenii de specialitate cunoscuţi numai unor cercuri restrînse de vorbitori, criteriul uzajului unanim prezintă un grad de pertinenţă redus, întrucît nu face decît să împartă lexicul în două mari subclase: de o parte — cuvintele ştiute şi folosite de toată lumea, de alta — restul vocabularului. O gradare mai fină reclamă utilizarea unor parametri susceptibili de mai multe valori, condiţie îndeplinită de primele trei criterii menţionate, dar nu şi de acesta din urmă.

Atît selecţia operată cît şi unele afirmaţii făcute cu prilejul comentării şi justificării acesteia evidenţiază cu claritate grija lui Al. Graur de a păstra o justă cumpănă între criterii[1], ferindu-se de a acorda vreunuia dintre ele o prioritate absolută. Instanţa care decidea acceptarea cuvintelor în listele FPL rămînea, pînă la urmă, discernămîntul lingvistic, întemeiat deopotrivă pe intuiţie şi pe bogata experienţă de lexicolog a autorului.

Dat fiind că intenţia noastră este să procedăm altminteri, şi anume să dăm prioritate aplicării riguroase şi imparţiale a unor criterii cuantificabile, cu scopul de a reduce la minimum marja de subiectivitate a cercetării, analiza întreprinderii lui Al. Graur este pentru

---

[1] „Criteriile generale de care m-am folosit în alegerea cuvintelor sînt cele expuse în introducere, dar aplicarea lor nu s-a făcut în chip mecanic" (*op. cit.*, pag. 39)



Scanned with OKEN Scanner

noi de cel mai mare interes, deoarece aceasta constituie atît termenul de referinţă, cît şi modelul exemplar al unui demers de sens contrar, în care flerul precumpăneşte asupra rigorii metodologice şi „simţul limbii" trece înaintea ingratei contabilităţi statistice[1].

## 4. Valoarea şi limitele fondului principal lexical în caracterizarea genealogică a vocabularului românesc

Conştient de dificultatea trasării unei frontiere nete între FPL şi masa vocabularului, Al. Graur a preferat arbitrariului unei soluţii univoce o rezolvare mai elastică a problemei, propunînd trei variante concentrice de vocabular fundamental:

— un cerc interior, pe care în cele ce urmează îl vom desemna prin $F_1$, conţinînd 965 de cuvinte, „cele mai sigure" sub raportul apartenenţei la fondul principal lexical;

— un cerc mijlociu, $F_2$, care include toate cuvintele din $F_1$ şi, în plus, alte 216 unităţi cu şanse ceva mai scăzute de a face parte din FPL;

— un cerc exterior, $F_3$, care mai adaugă la $F_2$ încă 238 de cuvinte a căror apartenenţă la FPL este şi mai puţin probabilă.

Odată stabilite listele de cuvinte corespunzătoare fiecăreia dintre cele trei variante, Al. Graur a analizat repartiţia etimologică a materialului respectiv, prilej de a observa că distribuţia procentuală pe origini variază în funcţie de extensiunea „cercului" considerat.

Astfel, cu cît numărul total de cuvinte este mai redus, cu atît creşte ponderea elementului latin: în $F_3$ cuvintele moştenite din latineşte reprezintă cca 58%, în $F_2$ — 62%, iar în $F_1$ — 66%. Pe de altă parte, din numărătoarea efectuată de D. Macrea[2] ştim că în Dicţionarul

---

[1] Se pare, de altfel, că, la fel cum nu a apelat la statistici privind frecvenţa cuvintelor, AL. Graur nu a socotit necesar să ordoneze în prealabil materialul lexical nici potrivit celorlalte criterii, ci s-a mulţumit în aprecierea lor cu elementele pe care i le oferea memoria. Aşa se explică, de pildă, faptul că a putut cita printre termenii care „nu au deloc derivate, dacă facem abstracţie de diminutive şi participii" (op.cit. pag. 24) cuvîntul *gură,* cu toate că în orice dicţionar al limbii române ar fi putut găsi măcar cuvintele *gures* şi *guraliv,* ambele nici diminutive, nici participii.

[2] D. Macrea: *Compoziţia istorică si tendinţele actuale de dezvoltare ale vocabularului românesc,* Cercetări de lingvistică, anul V, nr. 1—2, 1960, Cluj.

Limbii Române Moderne (notat de aici încolo prin DLRM), lucrare contemporană cu cea a lui Al. Graur (1958) şi, ca atare, reflectînd practic acelaşi stadiu istoric de evoluţie a limbii române, cuvintele latineşti moştenite totalizează numai 3,7% din întregul inventar. Aşadar, cu cît înaintăm către centrul vocabularului, densitatea elementului latin creşte, observaţie valabilă şi pentru alte influenţe mai vechi (substratul, împrumuturile vechi slave).

Cu totul altfel stau lucrurile cu întinsa categorie a cuvintelor de provenienţă recentă: proporţia acestora este cu atît mai mare cu cît orizontul lexical considerat e mai larg. Grupate în general sub eticheta „origine franceză", deşi nu întotdeauna provenienţa lor nemijlocit galică este uşor de stabilit, neologismele ocupă în DLRM (după D. Macrea) 25,7% din totalul cuvintelor, în timp ce la Al. Graur, care preferă, tocmai din pricina incertitudinilor de atribuire să le numească „internaţionale", ele reprezintă 6,7% din $F_3$, 4,57% din $F_2$ şi numai 2,69% din $F_1$.

Nimic surprinzător în constatările de mai sus. Ele nu fac decît să confirme adevărul cunoscut că elementul latin moştenit formează nucleul reprezentativ al lexicului românesc, în timp ce împrumuturile recente ocupă o poziţie periferică, penetraţia lor în miezul intim al limbii rămînînd modestă, în ciuda prezenţei lor masive în ansamblul vocabularului.

În măsura în care acordă un loc privilegiat componentei latine şi, în general, împrumuturilor mai vechi, statistica lui Al. Graur pare într-adevăr mai relevantă pentru profilul lexical al limbii române decît despuierile exhaustive ale unor dicţionare, precum cele efectuate de D. Macrea pentru Dicţionarul Enciclopedic Ilustrat „Cartea Românească" al lui I.A. Candrea[1] şi Dicţionarul Limbii Române Moderne. Nu vom subscrie însă la concluzia lui Al. Graur că „împărţirea brută a cuvintelor pe origini [...] fără să se ţină seamă de importanţa lor, duce la rezultate neconcludente şi neinteresante" (*op.cit.*, p. 60). Sperăm că, din analiza pe care o vom întreprinde în paragrafele următoare, să se desprindă încheierea că dicţionarul, fie acesta DLRM sau oricare altă lucrare lexicografică bine întocmită, reprezintă o secţiune prin vocabularul limbii române cu nimic mai puţin „concludentă şi interesantă" decît fondul principal lexical. Pînă atunci, să observăm numai că, importanţa cuvintelor fiind o mărime gradual variată, dependentă ea însăşi de o sumedenie de alţi parametri

[1] D. Macrea: *„Circulaţia cuvintelor în limba română"*, Transilvania, anul LXXIII, nr. 4, 1942.



Scanned with OKEN Scanner

lingvistici, deziderатul de a opera exclusiv cu cuvinte egal de importante este practic irealizabil. Dacă e perfect adevărat că nu toate unităţile din dicţionar au în limbă aceeaşi însemnătate, lucrul rămîne valabil pentru oricare altă mulţime de cuvinte. Din acest punct de vedere nici listele elaborate de Al. Graur nu fac şi nu puteau face excepţie[1]. Orice s-ar spune, verbul *a fi* ocupă în limba română o poziţie mai însemnată decît substantivul *valmă,* deşi ambele figurează la Al. Graur în cercul interior $F_1$ al fondului principal. Prezenţa lui *valmă* în FPL se justifică într-adevăr prin mulţimea derivatelor *(valmeş, vălmăşag, devălmăşie, învălmăşi* etc.) şi existenţa cîtorva expresii ca *de-a valma* şi *a da (intra) valma,* dar de aici şi pînă la a-l pune pe acelaşi plan cu multifuncţionalul verb al fiinţării este o distanţă considerabilă. *Toute proportion gardée,* FPL ne apare şi el, nu altfel decît dicţionarul, ca un teren al inegalităţii şi privilegiilor individuale. Cuvintele alese de Al. Graur sînt într-adevăr mai importante, de regulă, decît cele rămase în afara selecţiei sale, însă în nici un caz *egal de importante* şi oricînd s-ar putea delimita în sînul cercului interior $F_1$ alte şi alte submulţimi concentrice de cuvinte de însemnătate din ce în ce mai mare, mergînd pînă la un ipotetic cel mai important cuvînt românesc, a cărui identitate rămîne de precizat.

Variaţia, constatată de Al. Graur, a ponderii relative a diferitelor provenienţe în funcţie de extensiunea vocabularului considerat ridică în mod firesc următoarea întrebare: presupunînd drept posibilă ordonarea după importanţă a întregului lexic românesc, ce repartiţie pe origini vom obţine dacă luăm în considerare numai primele 10, 20, ..., 100, 1000, 10 000 sau, în general, *n* cuvinte, adică dacă optăm pentru un vocabular fundamental dimensionat fie cu mai multă generozitate, fie mai parcimonios decît cele trei variante propuse pentru fondul principal lexical?

Răspunsul la această întrebare este de o însemnătate capitală pentru caracterizarea genealogică a vocabularului românesc, deoarece numai el poate oferi o imagine globală a raporturilor dintre diferitele componente etimologice ce stau la baza limbii române. Orice descriere a situaţiei actuale a vocabularului, întemeiată exclusiv pe analiza distribuţiei pe origini a unei anumite colecţii de cuvinte, indiferent

---

[1] O recunoaşte însuşi autorul citat cînd observă: „Există prin urmare şi în sînul fondului principal o ierarhie, nu toate cuvintele sînt de aceeaşi tărie: unele se pregătesc să treacă înapoi în masa vocabularului, altele de-abia au intrat în fondul principal şi nu sînt bine consolidate, altele sînt solid instalate şi nu vor putea fi mişcate din locul lor multă vreme de aici înainte" (*op. cit.,* p. 43).



Scanned with OKEN Scanner

cum ar fi stabilită această colecţie, nu furnizează decît adevăruri parţiale, care, absolutizate în virtutea respectului pe care îl impune, mai întotdeauna, o contabilitate exactă, riscă să conducă la concluzii eronate.

Am văzut că fondul latin moştenit, de departe precumpănitor în vocabularul fundamental considerat de Al. Graur, devine, la nivelul dicţionarelor inventariate de D. Macrea, o componentă cantitativ modestă a lexicului românesc, strivită de masa neologismelor de sorginte franceză. Faţă de concluziile „uluitoare" (termenul îi aparţine lui Al. Graur[1]) la care ajunsese D. Macrea în urma numărării cuvintelor de diferite proveniențe din *Dicţionarul limbii române din trecut şi de astăzi*[2] al lui I. A. Candrea, şi anume că ponderea împrumuturilor din franceză (cca 30%) este superioară celei a elementului latin (cca 20%), Al. Graur exclama[3]: „Aproape că nici nu mai există o limbă română!" Lucrul părîndu-i-se de necrezut, lingvistul a căutat explicaţia situaţiei în aceea că „Dicţionarul cosmopolit al lui Candrea trece ca franţuzeşti fel de fel de cuvinte de altă origine" (ceea ce este parţial adevărat, dar nu într-o proporţie de natură să schimbe radical datele problemei) şi că multe franţuzisme din Dicţionar sînt termeni de jargon „pe care oamenii muncii nu le cunosc, nu le-au folosit şi nu le vor folosi niciodată", conchizînd în final: „N-am făcut socoteala exactă, dar sînt convins că elementele sigur franceze din limba noastră nu trec de 3 sau 4 la sută".

N-au trecut decît 6 ani şi această estimare greu de susţinut s-a văzut dezminţită de o nouă statistică a lui D. Macrea, bazată de această dată pe o lucrare academică deloc „cosmopolită": *„Dicţionarul limbii române moderne"* din 1958. Contrar aşteptărilor lui Al. Graur, raportul dintre elementul neologic francez şi fondul latin s-a modificat în favoarea celui dintîi. Acesta a ajuns să reprezinte peste 38% din total, în timp ce cuvintele de origine latină (inclusiv derivatele) şi-au menţinut cota de 20%. Cunoscut find ritmul rapid de asimilare a neologismelor ce caracterizează actuala etapă de evoluţie a limbii române, este de aşteptat ca, în lucrările lexicografice ulterioare, decalajul să fie şi mai spectaculos.

Ameninţă aceasta cu ceva statutul limbii noastre de continuatoare a latinei dunărene? Evident, cîtuşi de puţin. Şi nu neapărat numai

---

[1] Al. Graur: *Încercare* ... pag. 34

[2] Lucrarea constituie partea I a *Dicţionarului Enciclopedic Ilustrat* „Cartea Românească" din anul 1931

[3] *Op. cit.* pag. 35



Scanned with OKEN Scanner

fiindcă limba care are un impact atît de considerabil asupra românei este tot un idiom romanic, ci mai ales, şi aici Al. Graur are perfectă dreptate, fiindcă de-a lungul timpului, vocabularul fundamental şi-a conservat componenta preponderent latină, dovedindu-se foarte stabil şi puţin permeabil pentru împrumuturile recente. Analiza provenienţei cuvintelor din fondul principal lexical ne dă o idee despre acest primat al elementului latin, dar nu şi imaginea sa completă, avînd în vedere că ea se referă doar la un caz particular şi anume la repartiţia pe origini a celor mai importante 1181 de cuvinte (eventual 965 sau 1419 în cazul „cercurilor" înscris ori circumscris). Dacă, în funcţie de extinderea FPL, componenta lexicală latină reprezintă 58%, 62% şi, respectiv, 66% din numărul total al cuvintelor, ne putem aştepta ca, într-o variantă şi mai restrînsă de vocabular fundamental, ea să ajungă la 70% sau 80%. În aprecierea ponderii bazei latine a limbii române, partizanii fervenţi ai romanităţii noastre vor fi astfel tentaţi să se oprească la un fond mai redus de termeni esenţiali, pe cînd adversarii continuităţii vor prefera să ia în considerare un vocabular cît mai bogat, în care elementele moştenite să se piardă în oceanul împrumuturilor. După cum se vede, este loc aici pentru justificarea *cu cifre exacte* (!) a oricărei idei preconcepute.

Evitarea acestui paradox nu o vedem posibilă decît prin găsirea unor relaţii generale care să permită precizarea distribuţiei numerice pe origini a cuvintelor din mulţimile de tip „vocabular fundamental" de orice extensie. Ponderea elementului latin va fi exprimată atunci nu printr-o unică valoare procentuală, ci printr-o ecuaţie definită pe întregul domeniu de existenţă al lexicului românesc. Alte ecuaţii similare vor descrie aporturile slav, maghiar, turcesc, grecesc, german etc. Ordinea de importanţă a acestor componente în constituirea faciesului lexical al limbii române se va obţine în acest caz din analiza comparativă a coeficienţilor ecuaţiilor respective şi, în măsura în care acestea dau seamă de întreaga complexitate a stratificării după importanţă a cuvintelor, ierarhia obţinută va fi inatacabilă.

Obiectivul este, cum bine se observă, deosebit de ambiţios. Avînd în vedere că abordarea frontală a problemei ridică mari dificultăţi, vom încerca în prealabil să tatonăm terenul, pornind de la cazurile particulare cunoscute. Examinarea structurii celor trei variante de fond principal lexical propuse de Al. Graur ne poate fi în această privinţă de un real folos. În ciuda manierei intuitive, mai puţin riguroase, de alcătuire a lor, sau poate tocmai de aceea, aceste liste de termeni fundamentali oferă sugestii interesante privind tipologia analitică a relaţiilor căutate.



Scanned with OKEN Scanner

## 5. Repartiţia după provenienţă a cuvintelor din fondul principal lexical

Gruparea pe origini a cuvintelor din FPL l-a condus pe Al. Graur la inventarul[1] din tabelul 5.1.

*Tabelul 5.1.*

| | l | r(l) | s | r(s) | n. | i. | m. | a. | t. | g. | f. | o. | e. | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $F_1$ | 645 | 19 | 192 | 1 | 37 | 24 | 18 | 9 | 6 | 11 | 1 | 2 | — | 965 |
| $F_2$ | 731 | 25 | 245 | 2 | 58 | 46 | 24 | 18 | 10 | 13 | 6 | 2 | 1 | 1 181 |
| $F_3$ | 827 | 30 | 305 | 2 | 68 | 80 | 33 | 23 | 14 | 18 | 12 | 6 | 1 | 1 419 |

Semnificaţia abrevierilor este următoarea: a = autohton, e = englez, f = francez, g = grecesc, i = internaţional, l = latin, m = maghiar, n = necunoscut, o = onomatopeic, r(l) = format în româneşte din elemente latine, r(s) = idem din elemente slave, s = slav, t = turcesc.

Privit în lumina scopului urmărit de noi, acest inventar se dovedeşte nesatisfăcător din următoarele motive:

1. Tratamentul nepertinent aplicat de Al. Graur cuvintelor de origine slavă. Într-adevăr, din pricini asupra cărora nu este cazul să ne oprim aici, Al. Graur taxează nediferenţiat drept *slave* cuvinte care provin atît din slava veche cît şi, prin împrumut mai recent, din limbi neoslave ca bulgara, sîrbocroata, rusa, ucrainiana, polona. Procedeul viciază raporturile cantitative cu elementul romanic, ce nu beneficiază de un regim similar, fondul latin moştenit nefiind, la rîndul său, sporit prin adăugirea împrumuturilor de sursă latină savantă şi romanică (italiană, franceză). Oricum, indiferent de considerentele care au stat la baza demersului lui Al. Graur, în ceea ce ne priveşte, disocierea pe idiomuri a contribuţiei slave este indispensabilă.

2. Rezolvarea prea tranşantă a unor etimologii neclare asupra cărora nu s-a ajuns încă la un consens. Frapează, astfel, de la prima vedere, absenţa totală din tabelul 5.1. a unei rubrici a elementelor de origine incertă (în DLRM, D. Macrea numărase, cu tot cu derivate,

[1] Valorile din tabelul 5.1. nu coincid întrutotul cu cele din *Încercare* ..., deoarece la reproducerea lui am ţinut seama de erata publicată de Al. Graur în lucrarea ulterioară *Tendinţele actuale ale limbii române*, Editura ştiinţifică, Bucureşti 1968 (v. nota de la pag. 274).



Scanned with OKEN Scanner

1349 de astfel de unităţi). Al. Graur preferă să omită acel *cf* care marchează de obicei în dicţionare existenţa unor dificultăţi de ordin fonetic ori semantic ce fac îndoielnică descendenţa cuvîntului din etimonul indicat. El include, de exemplu, în categoria unităţilor de origine maghiară cuvîntul *talpă,* insuficient explicat prin mag. *talp,* la care DLRM şi DEX trimit cu cf., în timp ce Maria Iliescu (SCL XXVIII, 1977, pag. 184) îl atribuie substratului, iar M. Sala şi I. Coteanu îl consideră de provenienţă probabil latină[1]. De acest mod de a proceda profită cu precădere tot elementul slav, căruia îi sînt repartizate fără rezerve 25 de cuvinte apropiate, dar nu deduse fără cusur din elemente vechi slave, bulgare, sîrbocroate, poloneze, ucrainiene, ruse.

3. Un procedeu de sens contrar este aplicat multor neologisme, cărora Al. Graur ezită să le precizeze originea, grupîndu-le laolaltă sub eticheta neutră şi prea puţin lămuritoare de „internaţional". Contribuţia hotărîtoare a limbii franceze în amplul proces de reromanizare a vocabularului românesc este astfel ocultată.

La toate acestea, trebuie adăugat un aspect ce nu constituie o carenţă decît în raport cu obiectivul pe care ni l-am propus. Intenţionăm, aşa cum am făcut-o de altfel şi în paragraful precedent, să întemeiem deducţiile ce urmează atît pe datele oferite de propunerea de fond principal a lui Al. Graur, cît şi pe statistica efectuată de D. Macrea pe baza DLRM, lucrare contemporană cu a precedentului şi care reflectă, ca atare, acelaşi stadiu de evoluţie a limbii. Or, inexistenţa la Al. Graur a unei clase a cuvintelor de origine nesigură sau tratarea nediferenţiată, egalizatoare a neologismelor, ca şi a elementelor de provenienţă slavă, fac imposibilă orice comparaţie între cele două bilanţuri. Pentru aducerea lor la acelaşi numitor, ne-am văzut nevoiţi să reluăm repartizarea pe origini a unităţilor din FPL, în conformitate cu indicaţiile etimologice din DLRM. Rezultatul acestei noi clasificări a materialului lexical din cele trei variante ale FPL este prezentat în tabelul 5.2.

*Tabelul 5.2.*

| | Latine | | Slave | | Necunoscute | | Rom. (latine) | | Franceze | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % |
| $F_1$ | 660 | 68,39 | 141 | 14,61 | 19 | 1,97 | 19 | 1,97 | 16 | 1,66 |
| $F_2$ | 751 | 63,59 | 179 | 15,16 | 30 | 2,54 | 25 | 2,12 | 39 | 3,31 |
| $F_3$ | 852 | 60,04 | 209 | 14,73 | 36 | 2,54 | 30 | 2,11 | 73 | 5,15 |

[1] *Etimologia şi limba română,* Editura Academiei R.S.R., pag. 105—107.



Scanned with OKEN Scanner

*(continuare) Tabelul 5.2*

| | Maghiare | | Nesigure | | Greceşti | | Bulgăreşti | | cf. Albaneze | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % |
| $F_1$ | 16 | 1,66 | 15 | 1,56 | 15 | 1,56 | 14 | 1,45 | 13 | 1,35 |
| $F_2$ | 23 | 1,95 | 20 | 1,70 | 15 | 1,27 | 17 | 1,44 | 26 | 2,20 |
| $F_3$ | 31 | 2,18 | 34 | 2,40 | 20 | 1,41 | 25 | 1,76 | 32 | 2,26 |

| | Bulg. şi Scr. | | Lat. savante | | Turceşti | | Rom. (slave) | | Ucr. şi rusă | | Scr. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % |
| $F_1$ | 12 | 1,24 | 7 | 0,73 | 5 | 0,52 | 3 | 0,31 | 2 | 0,21 | 2 | 0,21 |
| $F_2$ | 20 | 1,70 | 11 | 0,93 | 10 | 0,85 | 4 | 0,34 | 2 | 0,17 | 2 | 0,17 |
| $F_3$ | 27 | 1,90 | 17 | 1,20 | 13 | 0,92 | 4 | 0,28 | 2 | 0,14 | 4 | 0,28 |

| | Onomatopeice | | Germane | | Ruseşti | | Ucr. şi Poloneze | | Bulg., Scr. şi Pol. | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % |
| $F_1$ | 1 | 0,10 | 1 | 0,10 | 1 | 0,10 | 1 | 0,10 | 1 | 0,10 |
| $F_2$ | 1 | 0,08 | 1 | 0,08 | 1 | 0,08 | 1 | 0,08 | 1 | 0,08 |
| $F_3$ | 3 | 0,21 | 1 | 0,07 | 1 | 0,07 | 1 | 0,07 | 1 | 0,07 |

| | Bulg., Scr. şi Ucr. | | Poloneze | | Ucrainiene | | Total |
|---|---|---|---|---|---|---|---|
| | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. | % | Nr. cuv. % |
| $F_1$ | 1 | 0,10 | — | — | — | — | 965 |
| $F_2$ | 1 | 0,08 | 1 | 0,08 | — | — | 1 181 |
| $F_3$ | 1 | 0,07 | 1 | 0,07 | 1 | 0,07 | 1 419 |

În raport cu inventarul lui Al. Graur, practic numărul de provenienţe distincte s-a dublat. Dintre cele 13 rubrici verticale ale tabelului 5.1. a dispărut una singură (cea a elementelor de origine engleză, care cuprindea un singur cuvînt, *sport,* socotit în DLRM drept pătruns pe filieră franceză[1], ceea ce pare mai plauzibil), în schimb s-au

[1] Potrivit unui principiu demult acceptat, se consideră drept sursă a unui cuvînt românesc idiomul care ni l-a împrumutat în mod nemijlocit, fără a se ţine seamă de obîrşia îndepărtată a cuvîntului şi de celelalte limbi prin care acesta a mai peregrinat înainte de a ne parveni. Astfel cuvintele *biserică, hîrtie, pită, metru,* deşi toate îşi au originea în greceşte, sînt în limba română de provenienţă, respectiv, latină, slavă veche, bulgară şi franceză.



Scanned with OKEN Scanner

mai adăugat tot pe atîtea, referitoare la influenţele bulgară, sîrbocroată, poloneză, ucraineană, italiană, latină savantă, la elementele de provenienţă incertă şi la unele cuvinte cu ascendenţă dublă (bulgară şi sîrbocroată, ucraineană şi rusă, ucraineană şi poloneză) şi chiar triplă (bulgară, sîrbocroată şi poloneză; bulgară, sîrbocroată şi ucraineană). Nu comentăm aceste diferenţe şi nici imaginea destul de diferită pe care ne-o oferă ele asupra provenienţei lexicului românesc fundamental, avînd în vedere că refacerea bilanţului lui Al. Graur nu a urmărit aici un scop polemic, ci doar pe acela de a pregăti date de bază cît mai „curate" pentru modelul pe care ne-am propus să-l construim.

## 6. Raporturile cantitative dintre provenienţe. „Ipoteza liniarităţii"

Deocamdată demersul nostru este de tip statistic. Or, din această perspectivă, influenţele lexicale care nu sînt ilustrate în FPL decît prin cîţiva reprezentanţi nu pot fi luate în considerare. De aceea ne vom limita, pentru moment, la contribuţiile ce au oferit fondului principal, în varianta sa mijlocie[1] $F_2$, cel puţin 10 cuvinte. În ordinea ponderii pe care o au în $F_2$, acestea sînt: elementul latin moştenit (751 de cuvinte), elementul slav (179), cel „francez" (39), necunoscut (30), autohton (26), maghiar (23), nesigur (20), bulgăresc & sîrbocroat (20), bulgăresc (17), grecesc (15), latin savant (11) şi turcesc (10). Am lăsat deoparte cuvintele formate în româneşte din elemente de origine latină. Prezenţa acestora contrazice de fapt regula de alcătuire a FPL, potrivit căreia trebuie să se acorde „fiecărei familii un singur loc"[2]. Prin includerea unei atare categorii se ajunge, de exemplu, ca derivatul *cerşi* să figureze alături de termenul-bază *cere,* ori *nebun* alături de *bun.* Acest mod de a proceda viciază calculul productivităţii lexicale a unităţilor din fondul principal (de care Al. Graur s-a ocupat în

[1] „... lista mijlocie, cea pe care o considerăm cea mai justă, este demnă de încredere cu privire la stabilirea procentelor pe limbile de origine" (Al. Graur *Încercare...* pag. 62).

[2] *Op. cit.* pag 43



Scanned with OKEN Scanner

aceeaşi lucrare[1], fiindcă, scindînd familiile de cuvinte în subfamilii, ne înfăţişează o imagine deformată a fecundităţii termenilor-matcă. Aşa, de pildă, din cele 34 de derivate pe care le are în DLRM, cuvîntul *bun* îl pierde nu numai pe *nebun*, ci şi cele 6 unităţi provenite din acesta *(nebunariţă, nebunatic, nebunească, nebunesc, nebuneşte, nebunie)*, ceea ce afectează serios poziţia sa în ierarhia celor mai productive cuvinte româneşti.

Dacă ne limităm însă în exclusivitate la unităţile primare, nederivate, se impune, pentru corectitudinea paralelei pe care o facem cu statistica lui D. Marcrea, să preluăm şi din aceasta numai cifrele privitoare la cuvintele-bază. Tabelul 6.1. prezintă sinoptic distribuţia pe origini a acestor unităţi atît în DLRM, cît şi în cele 3 „cercuri" ale fondului principal, pentru toate influenţele reţinute.

*Tabelul 6.1*

| | LATINE | SLAVE | „FRANCEZE" | NECUN. | CF. ALB. | MAGHIARE | NESIGURE | BULG. & SCR. |
|---|---|---|---|---|---|---|---|---|
| $F_1$ | 660 | 141 | 16 | 19 | 13 | 16 | 15 | 12 |
| $F_2$ | 751 | 179 | 39 | 30 | 26 | 23 | 20 | 20 |
| $F_3$ | 852 | 209 | 73 | 36 | 32 | 31 | 34 | 27 |
| DLRM | 1 849 | 1 133 | 12 770 | 1 440 | 81 | 443 | 526 | 212 |

| | BULG. | GREC. | LAT. SAV. | TURC. |
|---|---|---|---|---|
| $F_1$ | 14 | 15 | 7 | 5 |
| $F_2$ | 17 | 15 | 11 | 10 |
| $F_3$ | 25 | 20 | 17 | 13 |
| DLRM | 279 | 579 | 738 | 1 048 |

Plaja foarte întinsă de valori pe care se înşiră cifrele din tabelul 6.1, mergînd de la 5 (= numărul de cuvinte turceşti din $F_1$) pînă la 12 770 (= numărul de cuvinte „franceze" din DLRM), face dificilă reprezentarea lor grafică tale-quale. Inconvenientul se depăşeşte lesne prin figurarea acestor numere într-o scară logaritmică. Tabelul 6.2, întrutotul analog cu precedentul, diferă numai prin aceea că locul valorilor absolute l-au luat aici logaritmii naturali ai acestora.

[1] *Op. cit*, pag. 108



Scanned with OKEN Scanner

*Tabelul 6.2*

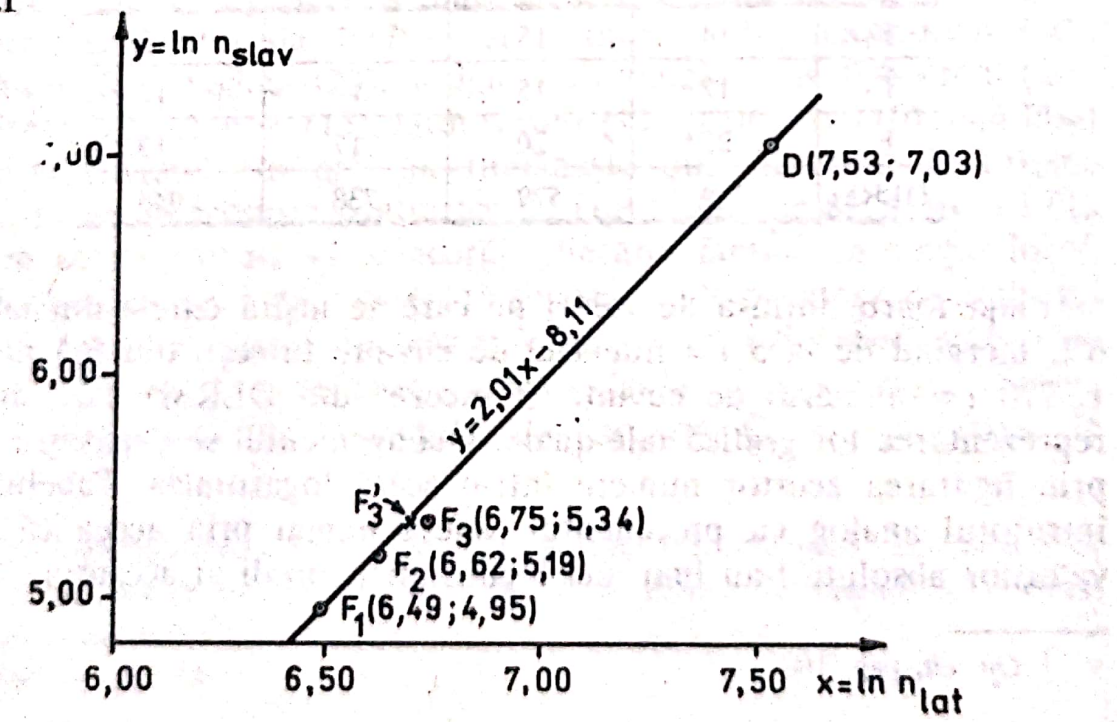
| | Latine | Slave | „Franc." | Necun. | cf. Alb. | Magh. | Nesig. | Bulg. & Scr. | Bulg. |
|---|---|---|---|---|---|---|---|---|---|
| $F_1$ | 6,49 | 4,95 | 2,77 | 2,94 | 2,56 | 2,77 | 2,71 | 2,48 | 2,64 |
| $F_2$ | 6,62 | 5,19 | 3,66 | 3,40 | 3,26 | 3,14 | 3,00 | 3,00 | 2,83 |
| $F_3$ | 6,75 | 5,34 | 4,29 | 3,58 | 3,47 | 3,43 | 3,53 | 3,30 | 3,22 |
| DLRM | 7,53 | 7,03 | 9,45 | 7,27 | 4,39 | 6,09 | 6,27 | 5,36 | 5,63 |

| | Grec. | Lat. Sav. | Turc. |
|---|---|---|---|
| $F_1$ | 2,71 | 1,95 | 1,61 |
| $F_2$ | 2,71 | 2,40 | 2,30 |
| $F_3$ | 3,00 | 2,83 | 2,56 |
| DLRM | 6,36 | 6,60 | 6,95 |

Putem obţine acum o imagine a raportului de forţe dintre oricare două idiomuri care au contribuit la formarea vocabularului limbii române, dacă reprezentăm într-un sistem cartezian, rectangular, de coordonate perechile de valori din tabelul 6.2 ce corespund aporturilor respective.

Figura 6.1 înfăţişează situaţia comparativă a elementului latin moştenit şi a celui vechi slav. Pe axa absciselor s-au înscris logaritmii naturali ai numărului de cuvinte de origine latină, iar pe cea a

FIG.6.1



ordonatelor — logaritmii numărului de cuvinte de origine slavă. Fiecărei mulţimi de cuvinte ce conţine elemente de ambele provenienţe îi corespunde un punct determinat din planul graficului. Astfel punctele $F_1$, $F_2$, $F_3$ şi D exprimă balanţa cantitativă dintre aportul latin şi cel slav în cele trei „cercuri" ale FPL şi în DLRM.

Se observă că trei dintre aceste puncte sînt practic coliniare (e vorba de $F_1$, $F_2$, şi D), în timp ce al patrulea ($F_3$) se abate întrucîtva de la dreapta pe care ele o determină. Ecuaţia dreptei este $y = 2{,}01\ x - 8{,}11$. Importanţa acestei ecuaţii constă în faptul că ea ne permite, desigur în măsura în care reflectă în mod corect realitatea lingvistică, să evaluăm contribuţia celor două surse lexicale în orice mulţime de cuvinte de tip vocabular fundamental[1], fie ea mai restrînsă sau mai extinsă decît fondul principal lexical. Astfel, dacă vrem să ştim, de exemplu, cîte cuvinte de obîrşie slavă va conţine un lexic de bază a cărui componentă latină se limitează la 100 de unităţi, nu ne rămîne decît să extragem dintr-un memorator matematic logaritmul numărului 100 (care este aproximativ 4,61) şi să introducem în ecuaţie valoarea respectivă:

$$x = 4{,}61 \qquad y = 2{,}01 \times 4{,}61 - 8{,}11 = 1{,}16$$

Pornind de la logaritmul $y = 1{,}16$, numărul de cuvinte de origine slavă se va obţine calculînd valoarea $e^{1{,}16}$, unde $e$ este baza de logaritmare ($\approx 2{,}72$). Rezultă $n_s = e^{1{,}16} \approx 3{,}2$, ceea ce înseamnă că într-un vocabular fundamental conţinînd 100 de cuvinte moştenite din latineşte nu vor figura decît 3 cuvinte de origine slavă! Iată un rezultat deloc banal, la care simpla cunoaştere a mereu citatei distribuţii procentuale calculate pentru FPL de Al. Graur nu ne-ar fi condus niciodată, căci ne-am fi aşteptat poate ca elementul slav să aibă o pondere de 3, 4 sau 5 ori mai mică decît cel latin, dar nu de 33 de ori!

Deocamdată însă liniaritatea relaţiei pe care am folosit-o constituie doar o ipoteză. Se verifică ea şi pentru alte perechi de influenţe?

Figurile 6.2 şi 6.3 prezintă, într-un mod asemănător celui descris anterior, raporturile elementelor „francez" şi, respectiv, turcesc cu

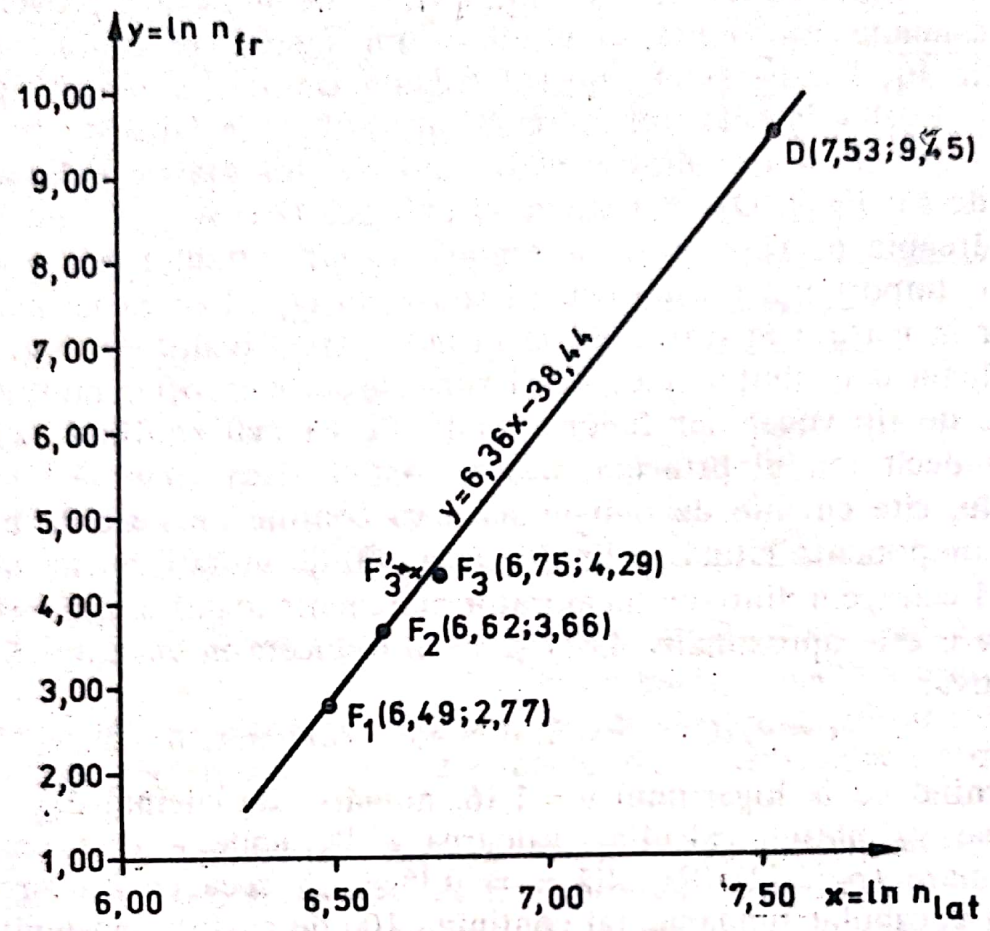
---

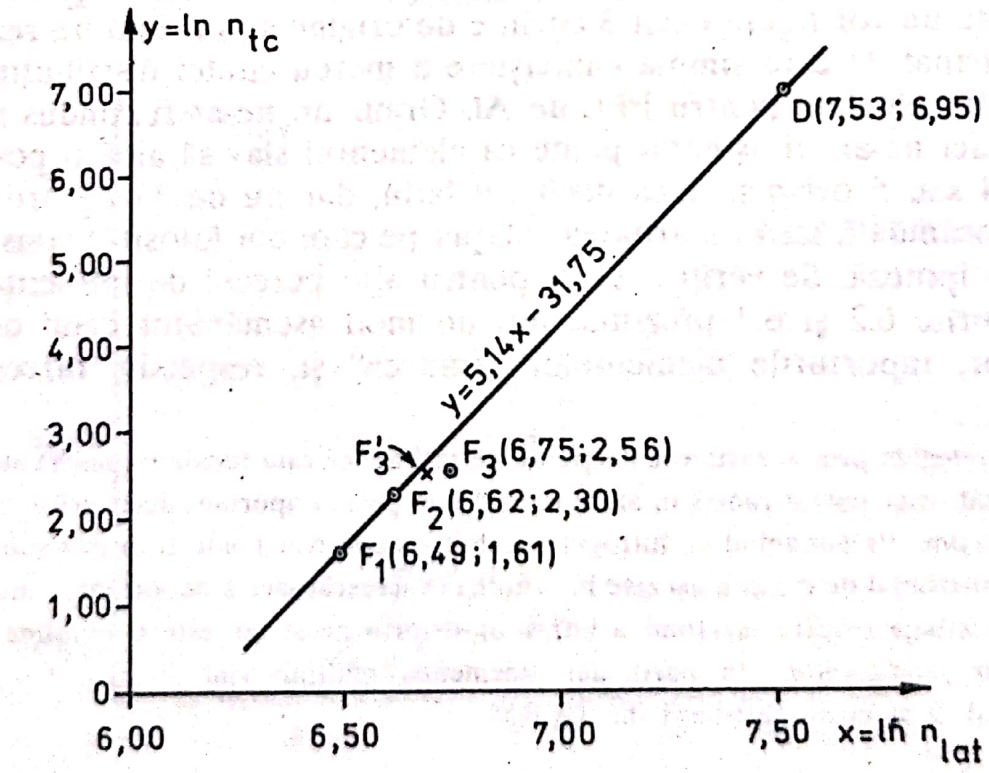
[1] Înţelegem prin aceasta o colecţie de unităţi nederivate (cuvinte-matcă) alcătuită astfel încît orice cuvînt rămas în afara ei este mai puţin important decît oricare cuvînt ce îi aparţine. Presupunînd că întregul vocabular nederivat poate fi dispus sub forma unui şir orizontal de cuvinte aşezate în ordinea descrescătoare a importanţei, mulţimea aflată la stînga oricărei secţiuni arbitrar alese prin acest şir este o *mulţime de tip vocabular fundamental*. În particular, asemenea mulţimi sînt $F_1$, $F_2$, $F_3$, dar şi ansamblul D al cuvintelor-bază din DLRM.



Scanned with OKEN Scanner

FIG. 6.2 şi 6.3






Scanned with OKEN Scanner

fondul lexical latin. În ambele cazuri vedem repetîndu-se situaţia dinainte: punctele $F_1$, $F_2$, şi D sînt aproape coliniare, iar $F_3$ se abate din nou şi *în acelaşi sens* ca pentru aportul slav de la dreapta trasată prin ele. „Ipoteza liniarităţii" cîştigă astfel un plus de credibilitate.

Cît priveşte poziţia deviantă a punctului $F_3$, ea s-ar putea explica acum printr-o oarecare supraestimare a ponderii elementului latin în „cercul exterior" al fondului principal. Într-adevăr, presupunînd că în $F_3$, nu s-ar găsi decît 812 cuvinte de origine latină în loc de 852 (ceea ce revine la o abatere de sub 5%), punctul $F_3$, s-ar deplasa în $F'_3$ apropiindu-se sensibil de dreapta caracteristică în toate cele trei grafice din figurile 6.1, 6.2 şi 6.3. Pentru respectarea întocmai a condiţiei de liniaritate a relaţiilor ar mai fi necesare doar ajustări minime ale numărului de cuvinte de origine slavă, „franceză" şi, respectiv, turcă din $F_3$. Calculul acestor corecţii este uşor de făcut, dacă se cunosc ecuaţiile dreptelor, ale căror expresii sînt:

$$y = 2{,}01\ x - 8{,}11 \qquad \text{pentru relaţia slav-latin} \qquad (6.1)$$

$$y = 6{,}36\ x - 38{,}44 \qquad \text{pentru relaţia „francez"-latin} \qquad (6.2)$$

$$y = 5{,}14\ x - 31{,}75 \qquad \text{pentru relaţia turcesc-latin} \qquad (6.3)$$

Considerînd pentru abscisa punctului $F_3$ valoarea $x = \ln 812 = 6{,}70$, găsim din cele trei ecuaţii valorile corespunzătoare ale ordonatelor şi, de aici, numărul de cuvinte slave, „franceze" şi turceşti care ar corespunde unui fond principal de 812 cuvinte latine moştenite. Rezultatul final al calculului este următorul:

$$n_{sl} = 212 \text{ cuvinte}, \qquad n_{fr} = 65 \text{ cuvinte}, \qquad n_{tc} = 15 \text{ cuvinte}$$

ceea ce înseamnă că pentru respectarea liniarităţii relaţiilor bilaterale dintre cele patru aporturi lexicale ar fi fost necesar ca în $F_3$ să figureze 212 cuvinte de origine slavă în loc de 209 considerate de Al. Graur, 65 de cuvinte „franceze", în loc de 73, şi 15 cuvinte turceşti în loc de 13. Ajustările ce ar trebui aduse selecţiei lui Al. Graur sînt, după cum se vede, modeste. Ne îndoim că cineva ar putea proba cu argumente irefutabile că „cercul" exterior al fondului principal trebuie să conţină cu necesitate 209 cuvinte slave şi nu 212, 73 cuvinte „franceze" şi nu 65 sau 13 cuvinte turceşti şi nu 15. Fireşte, contrariul este la fel de greu de demonstrat, însă nu aceasta contează, ci *posibilitatea* ca



Scanned with OKEN Scanner

lucrurile să stea şi aşa, suficientă deocamdată pentru a ne îngădui să afirmăm că „ipoteza liniarităţii" nu pare cu totul nerealistă[1].

Am pomenit mai sus de liniaritatea raporturilor bilaterale dintre toate influenţele considerate, deşi nu ne-am ocupat decît de relaţia diferitelor aporturi cu elementul latin, ceea ce poate părea o extrapolare abuzivă. Nu este însă aşa, cunoscută fiind tranzitivitatea raporturilor pe care le exprimă ecuaţiile de gradul întîi. Astfel, forma analitică a legăturii cantitative dintre aportul francez şi cel turcesc se obţine direct din ecuaţiile (6.2) şi (6.3) prin eliminarea parametrului care reprezintă componenta latină.

Notînd cu $n_{lat}$, $n_{fr}$ şi $n_{tc}$ numărul de cuvinte de origine latină şi, respectiv, „franceză" şi turcească, vom scrie:

$$ln\ n_{fr} = 6{,}36\ ln\ n_{lat} - 38{,}44$$

$$ln\ n_{tc} = 5{,}14\ ln\ n_{lat} - 31{,}75$$

Dacă explicităm din ambele relaţii parametrul $ln\ n_{lat}$, obţinem:

$$ln\ n_{lat} = \frac{ln\ n_{fr} + 38{,}44}{6{,}36} \tag{6.4}$$

$$ln\ n_{lat} = \frac{ln\ n_{tc} + 31{,}75}{5{,}14} \tag{6.5}$$

Prin egalarea expresiilor (6.4) şi (6.5), termenul referitor la aportul latin dispare şi se obţine o relaţie liniară între componenta „franceză" şi cea turcească. Forma acesteia, după rearanjarea convenabilă a termenilor este:

$$ln\ n_{fr} = 1{,}24\ ln\ n_{tc} + 0{,}83 \tag{6.6}$$

adică tot o ecuaţie de gradul întîi, care se mai poate scrie:

$$y = 1{,}24\,x + 0{,}83 \tag{6.7}$$

---

[1] Ghilimelele care încadrează sistematic în paragraful de faţă şi mai departe denumirea acestei conjecturi au dublul rost de a sublinia incertitudinile asociate momentan ipotezei respective, dar şi, mai cu seamă, faptul că nu e vorba de liniaritate decît în coordonate logaritmice, căci, în realitate, avem de a face cu ecuaţii exponenţiale, forma nelogaritmată a expresiilor 6.1 — 6.3 fiind:

$$n_{sl} = exp(2{,}01\ ln\ n_{lat} - 8{,}11)$$

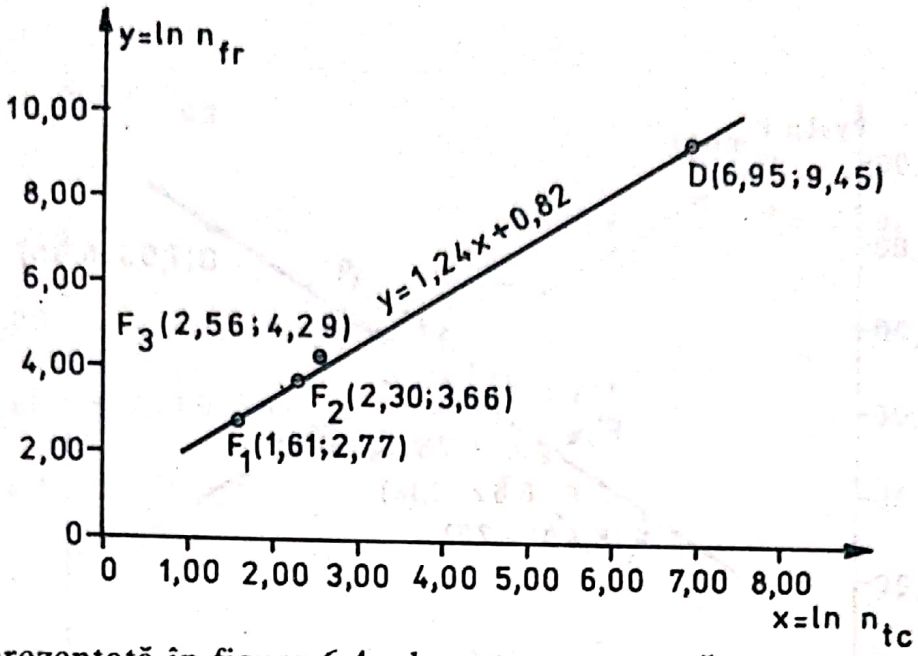
$$n_{fr} = exp(6{,}36\ ln\ n_{lat} - 38{,}44)$$

$$n_{tc} = exp(5{,}14\ ln\ n_{lat} - 31{,}75).$$



Scanned with OKEN Scanner

FIG.6.4




Reprezentată în figura 6.4., dreapta corespunzătoare trece, conform aşteptărilor, prin punctele $F_1$, $F_2$, D şi foarte aproape de poziţia punctului $F_3$.

Utilitatea observaţiei de mai sus constă în faptul că ea ne permite să restrîngem considerabil mulţimea relaţiilor de studiat. În loc să comparăm între ele toate cele 12 influenţe consemnate în tabelul 6.2, ceea ce ar da $C_{12}^2 = \frac{12 \cdot 11}{2} = 66$ de cupluri, va fi suficient să examinăm 11 situaţii, corespunzătoare raporturilor cu fondul latin ale restului de 11 contribuţii lexicale principale. Trei dintre ele au fost deja trecute în revistă. Cum se prezintă celelalte opt?

Componenta maghiară se supune şi ea unei legi liniare (fig. 6.5). Odată cu reducerea, preconizată anterior, a ponderii aportului latin în $F_3$ de la 852 la 812 cuvinte, punctul uşor deviant $F_3$ se deplasează şi el pe dreapta a cărei ecuaţie este:

$$y = 3{,}19x - 17{,}97 \qquad (6.8)$$

Avem deci în elementul maghiar un nou argument în favoarea „liniarităţii" şi, deopotrivă, a observaţiei potrivit căreia contribuţia latină la $F_3$ a fost uşor supraestimată.

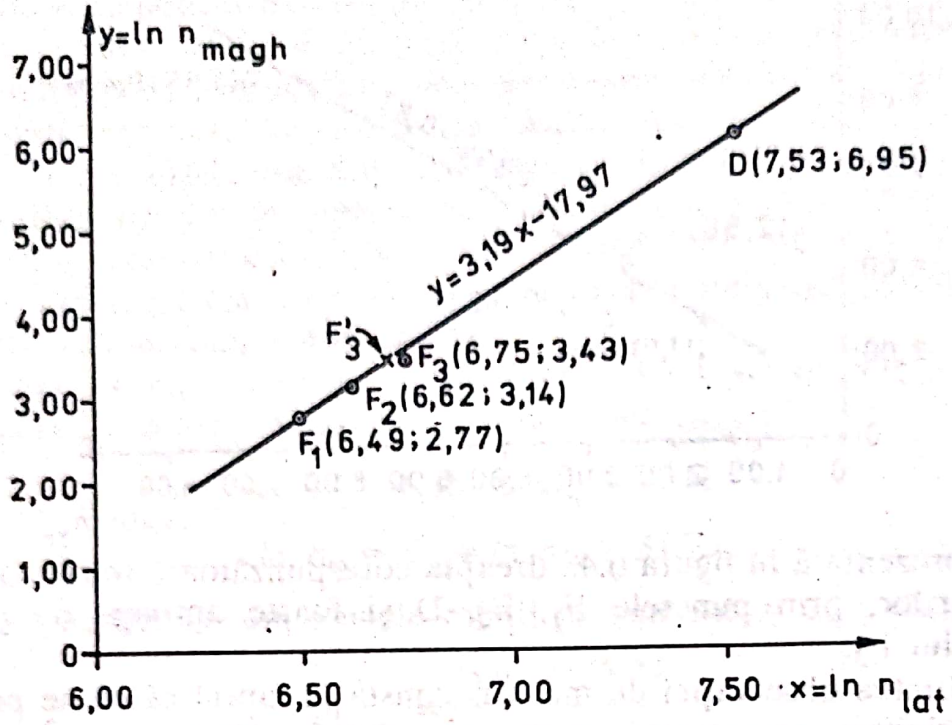
Încă alte patru componente istorice ale vocabularului sînt caracterizate prin coliniaritatea a cîte trei puncte semnificative, dar aceste puncte nu sînt întotdeauna aceleaşi.

În cazul cuvintelor comune cu albaneza, obişnuit şi, în general, justificat atribuite substratului, precum şi al împrumuturilor bulgăreşti şi mixte bulgaro-sîrbocroate, punctul nesituat pe dreaptă este $F_1$



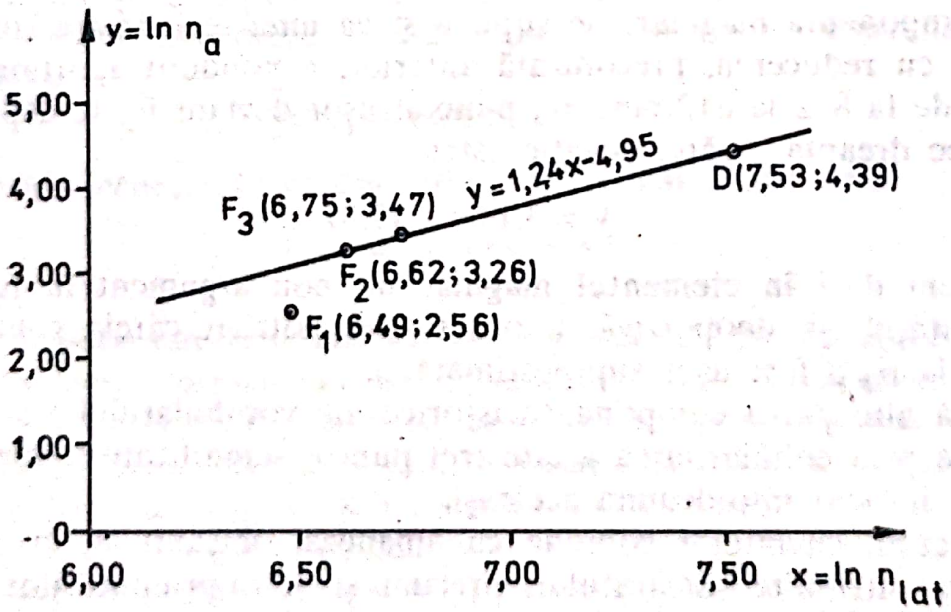
Scanned with OKEN Scanner

Fig. 6.5




(v. fig. 6.6, 6.7 şi 6.8). Cît de semnificativă este această abatere şi ce corecţii ar trebui aduse variantei celei mai restrictive a FPL pentru ca şi în aceste cazuri „liniaritatea" variaţiei cu fondul latin să fie respectată?

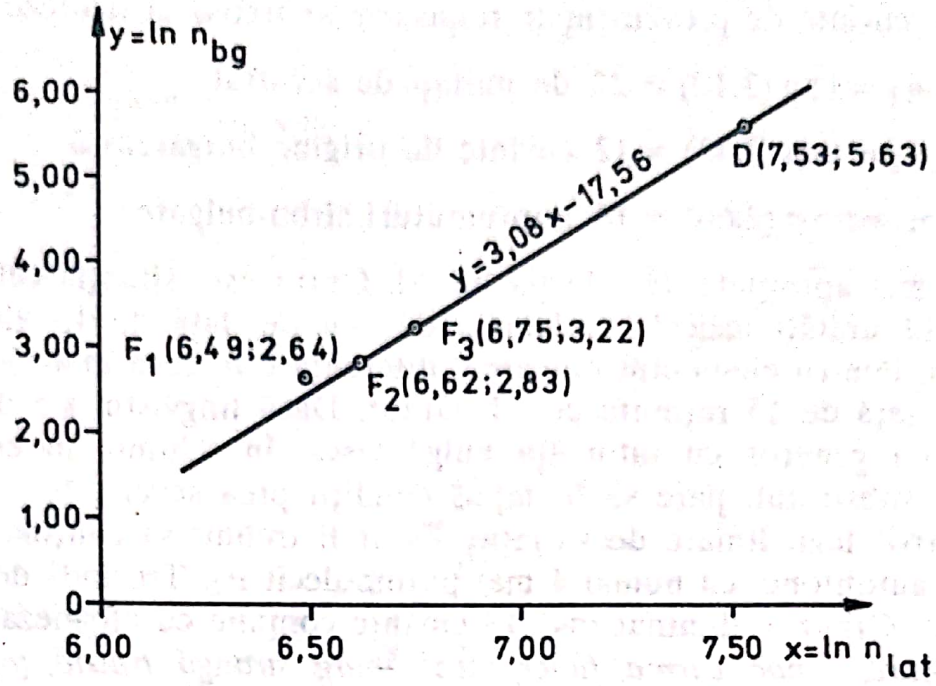
FIG. 6.6






Scanned with OKEN Scanner

FIG. 6.7



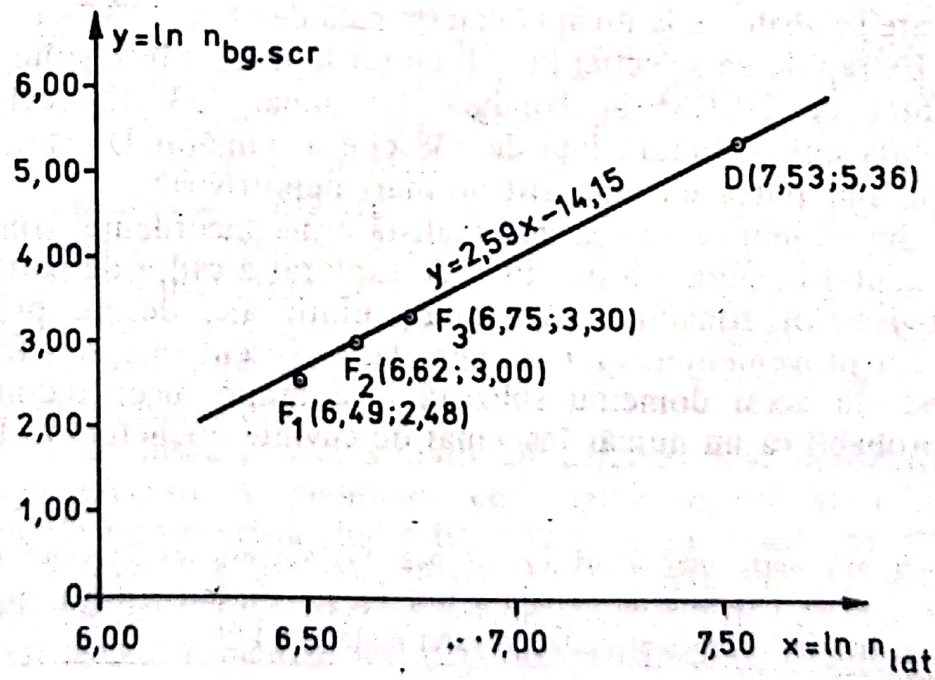

Avînd în vedere că $F_1$ cuprinde 660 de cuvinte latine moştenite, abscisa punctului $F_1$ este $x_1 = ln\ 660 \approx 6{,}49$. Introducînd această valoare în cele trei ecuaţii ale dreptelor din figurile 6.6, 6.7 şi 6.8, care sînt:

$y = 1{,}24x - 4{,}95$ pentru substrat (6.9)

$y = 3{,}08x - 17{,}56$ pentru elementul bulgăresc (6.10)

$y = 2{,}59x - 14{,}15$ pentru elementul bulgaro-sîrbocroat (6.11)

FIG.6.8






Scanned with OKEN Scanner

se obţin suucesiv $y_1 = 3{,}10$, $y_2 = 2{,}43$ şi $y_3 = 2{,}66$, de unde rezultă imediat cîte cuvinte de provenienţele respective ar trebui să figureze în $F_1$:

$n_1 = \exp(3{,}10) = 22$ de unităţi de substrat

$n_2 = \exp(2{,}43) = 12$ cuvinte de origine bulgărească

$n_3 = \exp(2{,}66) = 14$ împrumuturi sîrbo-bulgare

Cea mai apropiată de selecţia lui Al. Graur este situaţia celor din urmă (14 unităţi lexicale reclamate de ecuaţie, faţă de 13 alese de lingvist). Pentru elementul bulgăresc diferenţa este ceva mai mare: 12 cuvinte, faţă de 15 reţinute de Al. Graur. Dacă lingvistul s-a dovedit puţin mai generos cu influenţa bulgărească, în schimb, în ceea ce priveşte substratul, pare să fi impus condiţii prea severe de acces în $F_1$. Potrivit legii liniare de variaţie, $F_1$ ar fi trebuit să conţină 22 de cuvinte autohtone, cu numai 4 mai puţine decît $F_2$. Trecînd[1] de la $F_2$ la $F_1$ Al. Graur a eliminat însă 13 cuvinte comune cu albaneza: *abur, brad, cioară, copac, curma, fluier, mînz, murg, strungă, traistă, ţap, ţarc* şi *vatră*. Or, cu excepţia a cel mult 4—5 dintre ele, acestea ne par, şi am vrea să credem că nu sîntem părtinitori, fundamentale pentru bagajul lexical al oricărui vorbitor de limbă română. În cazul că impresia noastră ar fi justă, substratul s-ar putea număra şi el printre influenţele lingvistice care respectă legea liniară de variaţie cu ponderea elementului latin.

O poziţie aparte o ocupă împrumuturile de sursă latină savantă (fig. 6.9). De data aceasta, cele trei puncte coliniare sînt $F_1$, $F_2$ şi $F_3$, iar cel care se abate de la dreapta determinată de ele ($y = 3{,}38\,x - 19{,}99$) este D. Extrapolarea selecţiei lui Al. Graur la nivelul întregului corpus reprezentat de DLRM ar conduce la numai 223 de cuvinte de provenienţă latină literară faţă de 738 cîte a numărat D. Macrea. Ce explicaţie s-ar putea da unei atît de mari nepotriviri?

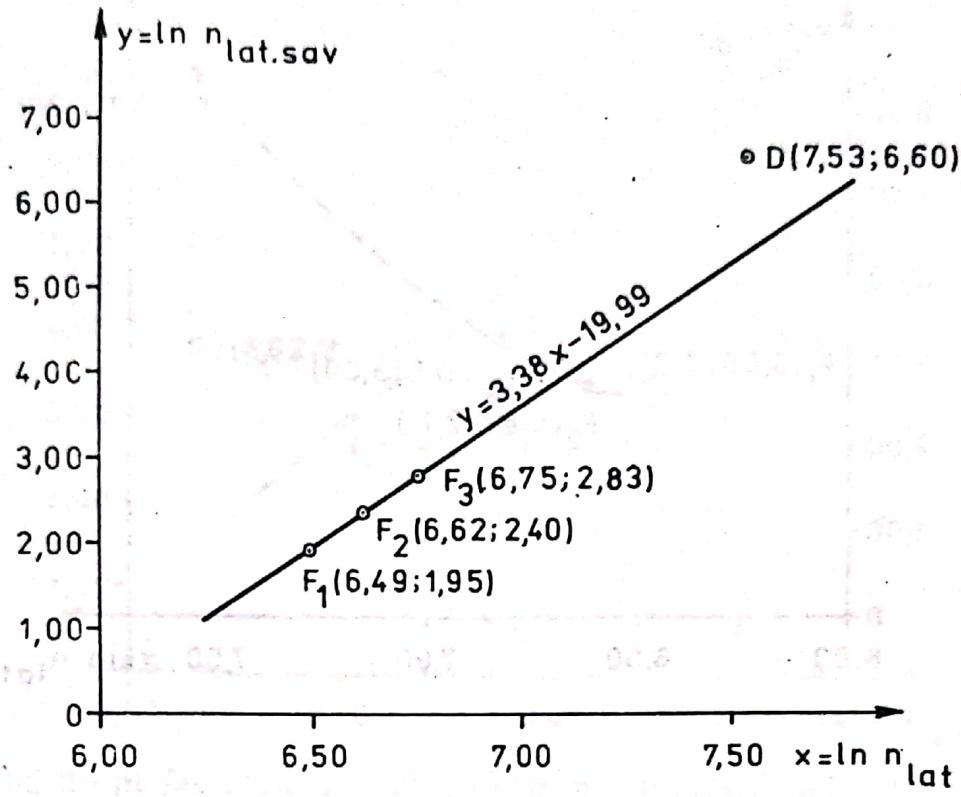
Socotim că motivarea cea mai realistă a neconcordanţei constatate ar fi de căutat în sfera atît de dificil de explorat a căilor de pătrundere a neologismelor romanice. Am mai amintit aici despre problema spinoasă a provenienţei unor termeni latini recent integraţi lexicului românesc. În acest domeniu subzistă încă multe incertitudini. Este foarte probabil ca un număr însemnat de cuvinte etichetate în DLRM

[1] Conform propriei mărturisiri (*op. cit.* pag. 39), succesiunea întocmirii de către Al. Graur a celor trei liste de cuvinte a fost $F_2$, $F_3$, $F_1$, într-o dublă mişcare de extindere şi restrîngere succesive ale variantei mijlocii a FPL.



Scanned with OKEN Scanner

FIG. 6.9




drept latine literare să fi pătruns în românește pe filiera unei limbi moderne, cu atît mai mult cu cît tradiția utilizării latinei ca mijloc de vehiculare a cunoștințelor științifice a fost la noi destul de firavă. Ni se pare de aceea plauzibil ca o parte însemnată dintre cele 738 de cuvinte latine savante din DLRM să ne fi parvenit prin mijlocirea altor limbi, fiind ca atare integrabile în principal clasei împrumuturilor „franceze". În raport cu numărul foarte mare al acestora din urmă (12.770), includerea a încă vreo 400—500 de cuvinte nu înseamnă decît o suplimentare cu vreo 3—4 procente, ceea ce nu schimbă esențial situația elementului „francez". De fapt, pentru valorile logaritmate cu care am operat, corecția ce trebuie adusă ar fi de numai 0,5—0,6%, ceea ce, la scara reprezentărilor noastre grafice, constituie o abatere de-a dreptul insesizabilă.

Rămîn de discutat trei situații cu caracter mai deosebit: cea a influenței grecești, a cuvintelor cu origine incertă și a celor de proveniență necunoscută (fig. 6.10, 6.11 și 6.12). Trăsătura comună a acestor componente ale vocabularului românesc este caracterul lor eteroclit. De fapt, toate trei se prezintă nu ca afluenți simpli ce au alimentat lexicul românesc într-o perioadă istoric determinată, ci mai



Scanned with OKEN Scanner

FIG. 6.10

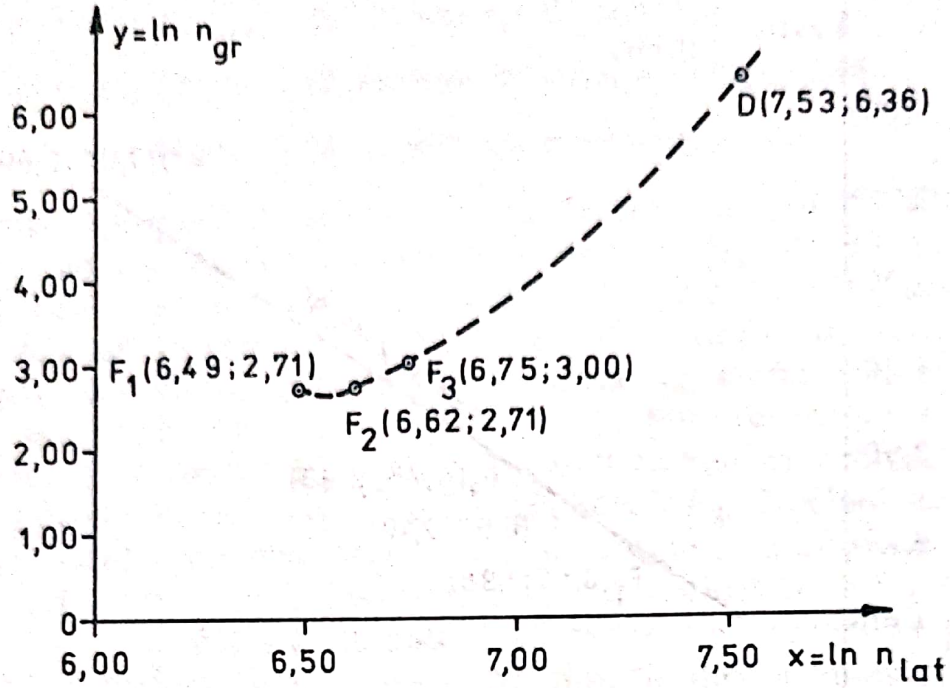
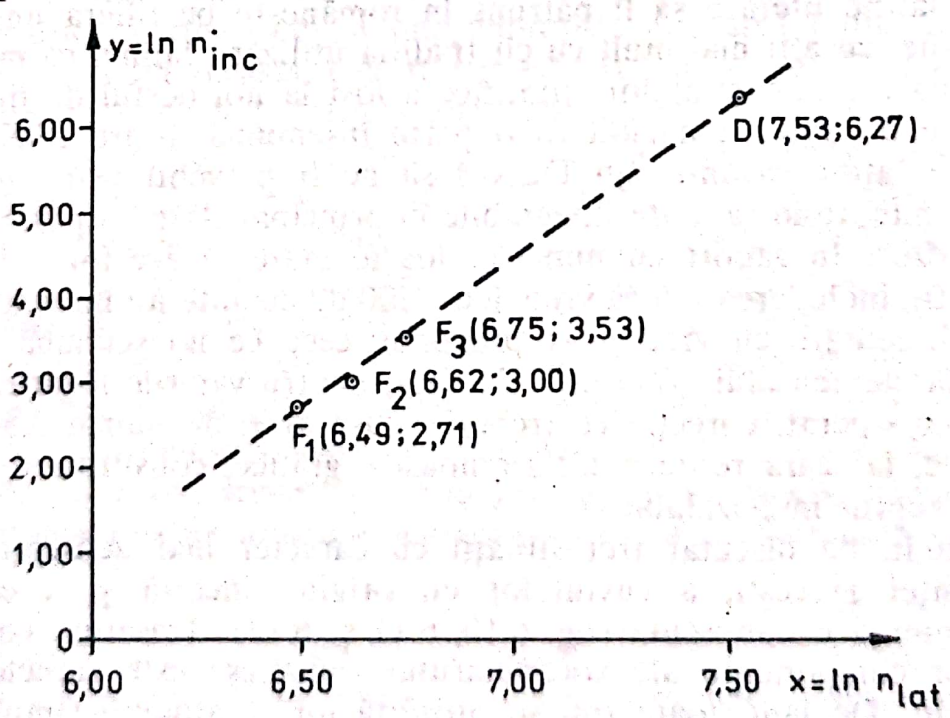


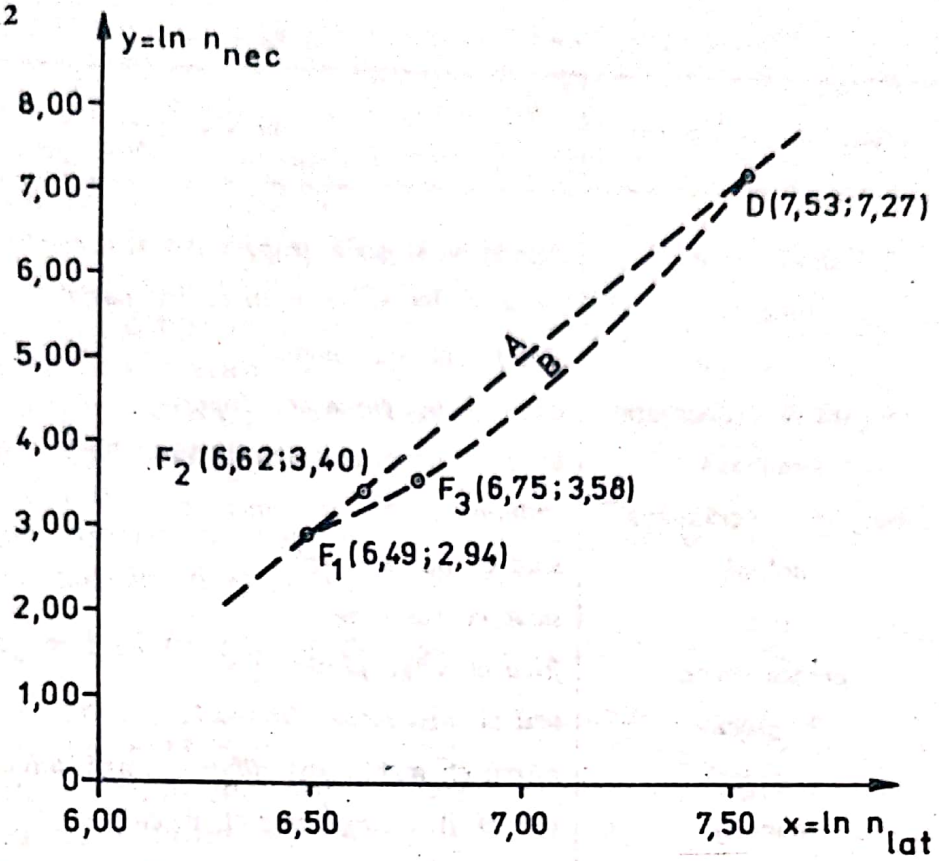

FIG.6.11






Scanned with OKEN Scanner

FIG. 6.12




degrabă drept fascicule de influenţe intricate acoperind un interval de timp greu de precizat (în orice caz foarte întins). Astfel, cuvintele zise „de origine necunoscută" sînt fără doar şi poate de obîrşii foarte variate, nici una dintre limbile cu care româna s-a aflat în contact de-a lungul timpului, începînd chiar cu substratul, nefiind în principiu exclusă ca posibilă sursă a acestor termeni. Că lucrurile stau aidoma şi în ceea ce priveşte elementele nesigure ne-o arată tabelul 6.3, din care reiese că unităţile cu etimologie incertă din FPL purced probabil din 12 izvoare diferite. Surprinzătoare poate părea numai asimilarea contribuţiei greceşti cu aceste influenţe hibride. De data aceasta, fiind vorba de o sursă lexicală unică, s-ar zice că ne aflăm într-un caz, similar celor anterior examinate, de aport omogen, circumscris temporal, cum erau influenţa slavă, maghiară sau turcească. Este însă bine cunoscut că aluviunile lexicale de origine grecească din limba română sînt rezultatul a cel puţin trei valuri succesive distincte. Există pe de o parte elenisme de dată latină[1], pe de altă parte împrumutur

[1] Pe cele general romanice *(carte, boteza)* lingviştii (DLRM nu face nici el excepţie) le consideră, pe bună dreptate, drept cuvinte latine moştenite, dar elenismele pătrunse pe cale orală în latina dunăreană *(proaspăt, frică)*, sînt socotite ca formînd un prim strat al influenţei greceşti asupra limbii române.



Scanned with OKEN Scanner

*Tabelul 6.3*

| Nr. crt. | Origine / înrudire probabilă | Exemple |
|---|---|---|
| 1 | slavă veche | *găsi* cf. v. sl *gasiti, ţeapă* cf. v. sl *cepati* |
| 2 | bulgară | *stejar* cf. bg. *stežer, trînti* cf. bg. *tărtja* |
| 3 | sîrbocroată | *sprinten* cf. scr. *spretan* |
| 4 | bulgară & sîrbocroată | *ciupi* cf. bg. *čupja* scr. *čupati* |
| 5 | ucraineană | *bubă* cf. ucr. *buba, slut* cf. ucr. *slutyj* |
| 6 | bulgară & ucraineană | *smîntînă* cf. bg. ucr. *smetana* |
| 7 | polonă | *sold* cf. pol. *szoldra* |
| 8 | rusă | *steag* cf. rus *stjag* |
| 9 | greaca veche | *fricâ* cf. v. gr. *phrike* |
| 10 | neogreacă | *ursi* cf. ngr. *orizo* |
| 11 | maghiară | *mereu* cf. mgh. *merö, tîlhar* cf. mgh. *tolvaj* |
| 12 | italiană | *ciot* cf. it. *ciotto, ciung* cf. it. *cionco* |

medio-greceşti din epoca bizantină[1] şi în fine cuvinte neogreceştii pătrunse cu precădere în epoca fanariotă. Separate în timp de intervale apreciabile şi reflectînd stadii diferite de evoluţie a limbii sursă, aceste trei straturi se prezintă ca tot atîtea influenţe diferite, la fel cum înrîurirea bulgărească sau rusească sînt altceva decît cea paleoslavă, iar neologismele romanice — decît fondul latin moştenit.

Cum se reflectă în aspectul grafic al relaţiei cu elementul latin caracterul compozit al alcătuirii unor mulţimi de cuvinte ca cele menţionate?

Pentru a putea răspunde la această întrebare, trebuie să reamintim observaţia că raporturile cantitative dintre diferitele componente istorice ale vocabularului românesc ne-au apărut drept liniare numai întrucît am apelat la o reprezentare în coordonate logaritmice, căci ele sînt de fapt de tip exponenţial.

[1] Şi în sînul acestora se disting două categorii: împrumuturi venite prin contact direct *(folos, prisos)* şi împrumuturi transmise prin intermediul limbii slave *(icoană, psalm)* sau bulgare *(pită, samar)*. Numai primele pot fi socotite elenisme autentice.



Scanned with OKEN Scanner

Dar se ştie că, în general, suma a două exponenţiale nu este şi ea o exponenţială, ceea ce înseamnă că prin cumularea a două influenţe distincte nu mai obţinem o influenţă reprezentabilă, în coordonate logaritmice, sub forma unei drepte. Ne putem convinge de acest adevăr printr-un exerciţiu simplu.

Să presupunem bunăoară, că, dintr-un motiv sau altul, contribuţia substratului şi cea a limbii maghiare la constituirea lexicului românesc ne-ar apărea drept indiscernabile. Tratate drept o influenţă unică, ar mai respecta ele atunci o lege liniară?

Să trasăm prin puncte variaţia cu elementul latin a valorilor cumulate corespunzătoare celor două surse. Notînd cu $n_{lat}$, $n_a$ şi $n_m$ numărul cuvintelor de origine latină şi, respectiv, autohtonă şi maghiară dintr-o aceeaşi variantă de vocabular fundamental, putem rescrie ecuaţiile (6.9) şi (6.8) sub forma:

$$ln\ n_a = 1{,}24\ ln\ n_{lat} - 4{,}95 \quad \text{şi} \quad ln\ n_m = 3{,}19\ ln\ n_{lat} - 17{,}97$$

Considerînd succesiv pe axa absciselor 7 puncte echidistante: $ln\ n_{lat}$ = 6,00; 6,25; 6,50; 6,75; 7,00; 7,25; 7,50, calculăm pe rînd valorile în aceste secţiuni ale ordonatelor $y_a$ şi $y_m$, de unde rezultă imediat numărul de cuvinte autohtone $n_a = exp\ (y_a)$ şi de origine maghiară $n_m = exp(y_m)$ corespunzătoare. Pentru fiecare valoare a argumentului $x$, adunăm $n_a$ cu $n_m$, iar suma rezultată o logaritmăm pentru a o putea reprezenta pe acelaşi grafic cu dreptele (6.8) şi (6.9). Succesiunea etapelor de calcul, precum şi valorile obţinute la fiecare pas, sînt sistematizate în tabelul 6.4.

*Tabelul 6.4*

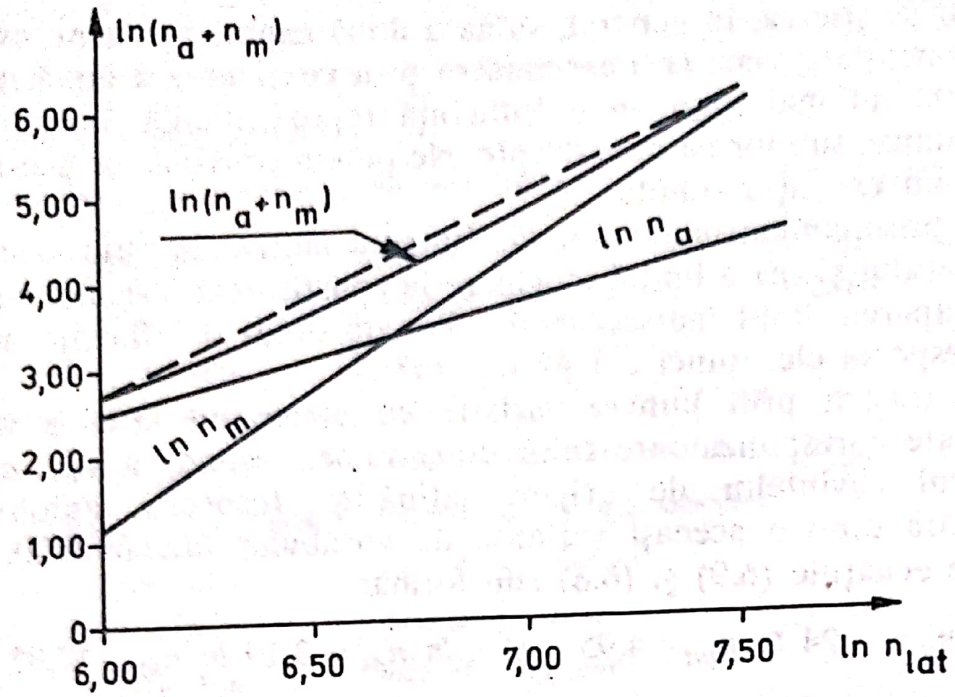
| $ln\ n_{lat}$ | 6,00 | 6,25 | 6,50 | 6,75 | 7,00 | 7,25 | 7,50 |
|---|---|---|---|---|---|---|---|
| $ln\ n_a$ | 2,49 | 2,80 | 3,11 | 3,42 | 3,73 | 4,04 | 4,35 |
| $ln\ n_m$ | 1,17 | 1,97 | 2,76 | 3,56 | 4,36 | 5,16 | 5,96 |
| $n_a$ | 12 | 16 | 23 | 31 | 42 | 57 | 78 |
| $n_m$ | 3 | 7 | 16 | 35 | 78 | 174 | 388 |
| $n_a + n_m$ | 15 | 23 | 39 | 66 | 120 | 231 | 466 |
| $ln\ (n_a + n_m)$ | 2,71 | 3,16 | 3,66 | 4,19 | 4,79 | 5,44 | 6,14 |

Prin reprezentarea rezultatelor din ultima linie a tabelului se obţine curba din figura 6.13 cu convexitatea îndreptată în jos, fapt pus în evidenţă prin figurarea cu linie întreruptă a coardei interioare. Forma ei demonstrează prin contraexemplu că la cumularea unor influenţe



Scanned with OKEN Scanner

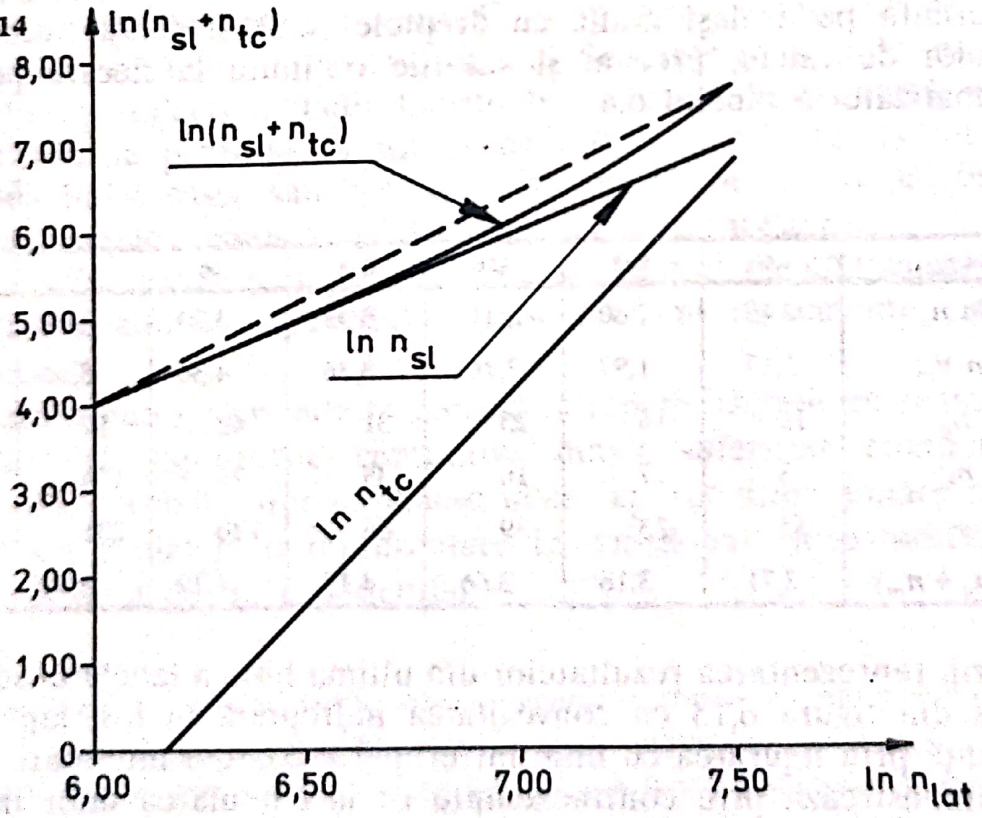
FIG.6.13




diferite caracterul exponenţial al corelaţiei cu elementul latin se pierde. La aceeaşi concluzie ne conduce (nu mai redăm aici calculele) şi însumarea forţată a aportului vechi slav cu cel turcesc (fig. 6.14).

În lumina acestor rezultate, ne aşteptăm ca aporturile eteroclite de felul împrumuturilor greceşti, al cuvintelor cu etimologie neclară sau

FIG. 6.14



al celor de proveniență necunoscută, să nu mai satisfacă „ipoteza liniarității". Conjectura se confirmă întrutotul în cazul dintîi: dispunerea punctelor corespunzătoare prezenței elementului grecesc în fondul principal și în DLRM (fig. 6.10) amintește de curbele largi, concave, din figurile 6.13 și 6.14. Pentru componentele „incert" și „necunoscut" alura scontată nu mai este însă la fel de evidentă. Unitățile de origine nesigură se apropie simțitor de o distribuție liniară (fig. 6.11), ceea ce s-ar putea explica fie prin efectul de compensare datorat însumării unui număr mare de aporturi lexicale diferite (demonstrația noastră avea în vedere suprapunerea a numai două influențe distincte), fie prin predominanța absolută a cuvintelor de o anumită origine sau vechime[1] care imprimă ansamblului un comportament unitar. În schimb, așezarea punctelor din cazul cuvintelor de origine necunoscută (fig. 6.12) nu ne permite să emitem nici o supoziție privind relația cu fondul latin, care ar putea îmbrăca tot atît de bine forma dreptei A (cu $F_3$ ca punct „deviant"), ca și pe cea a curbei B (în care caz $F_2$ ar fi punctul ce se abate de la regulă).

Tot ceea ce putem reține pentru moment din analiza întreprinsă în acest paragraf este că, dacă studiul aporturilor etimologic omogene pare să sugereze existența între ele a unei legături de tip liniar, examinarea situației componentelor eterogene, constînd în mănunchiuri de influențe diferite, solidarizate grație unui liant exterior, artificial (care constă în cazul de față în ignoranța sau șovăiala diagnosticianului etimolog), nu este de natură să ofere argumente nici pro nici contra unei atari ipoteze.

## 7. Relativitatea raporturilor numerice dintre constituenții istorici ai lexicului. „Paradoxul lui A. de Cihac"

Ce urmări de ordin mai general ar putea avea, în cazul confirmării, sale, „ipoteza liniarității"? Principalele implicații privesc problema caracterizării genealogice a limbii române. Spre deosebire de abordarea

[1] Vom vedea mai departe că panta dreptelor caracteristice depinde direct de vechimea aporturilor lexicale respective și că, în consecință, mai multe influențe apropiate în timp vor putea conduce la o rezultantă liniară, ceea ce nu s-a întîmplat cu împrumuturile grecești, întinse de-a lungul unui interval mult prea îndelungat.



Scanned with OKEN Scanner

pur statistică, ce nu furniza decît o cunoaştere punctuală a raporturilor dintre diferitele componente istorice ale vocabularului, nepermiţînd generalizări, întrucît valabilitatea rezultatelor se limita la masa lexicală cercetată, o metodă analitică de felul celei discutate aici face posibilă exprimarea globală a acestor raporturi. Ea îngăduie clasificarea după importanţă a contribuţiei diferitelor limbi ce au jucat un rol în constituirea lexicului românesc, explicînd şi integrînd totodată imaginile parţiale oferite de cercetările anterioare şi subsumîndu-le unei viziuni de ansamblu care le încorporează şi le depăşeşte.

Faptul că „liniaritatea" corelaţiilor numerice subiacente nu are deocamdată decît statutul unei conjecturi oricînd amendabile ne împiedică să avansăm prea departe pe acest teren. O vom face mai tîrziu, cînd, anticipăm puţin, se vor fi adunat suficiente argumente pentru a putea vorbi nu de o ipoteză, ci de o lege lingvistică în toată puterea cuvîntului. Pentru moment, ne mărginim la examinarea unui aspect care, insuficient înţeles, poate da şi a dat efectiv naştere în trecut unor controverse aprinse. Este vorba de variabilitatea derutantă a repartiţiilor pe origini ale cuvintelor româneşti, constatate de diferiţi lingvişti cu prilejul încercării de a descrie în termeni statistici profilul etimologic al limbii române.

Presupunînd valabilă generalizarea efectuată în paragraful precedent a datelor rezultate din repartizarea pe origini a cuvintelor din fondul principal şi din Dicţionarul Limbii Române Moderne, contribuţiile principalelor influenţe lexicale la alcătuirea vocabularului românesc se exprimă analitic, precum am arătat, ca funcţii exponenţiale de componenta latină cu ajutorul următoarelor ecuaţii:

$$n_a = exp\,(1{,}24\ ln\ n_{lat} - 4{,}95) \text{ pentru substrat}$$

$$n_{sl} = exp\,(2{,}01\ ln\ n_{lat} - 8{,}11) \text{ pentru elementul slav vechi}$$

$$n_{bg.scr.} = exp\,(2{,}59\ ln\ n_{lat} - 14{,}15) \text{ pentru aportul mixt bulgaro-sîrbocroat}$$

$$n_{bg} = exp\,(3{,}08\ ln\ n_{lat} - 17{,}56) \text{ pentru elementul bulgar}$$

$$n_m = exp\,(3{,}19\ ln\ n_{lat} - 17{,}97) \text{ pentru elementul maghiar}$$

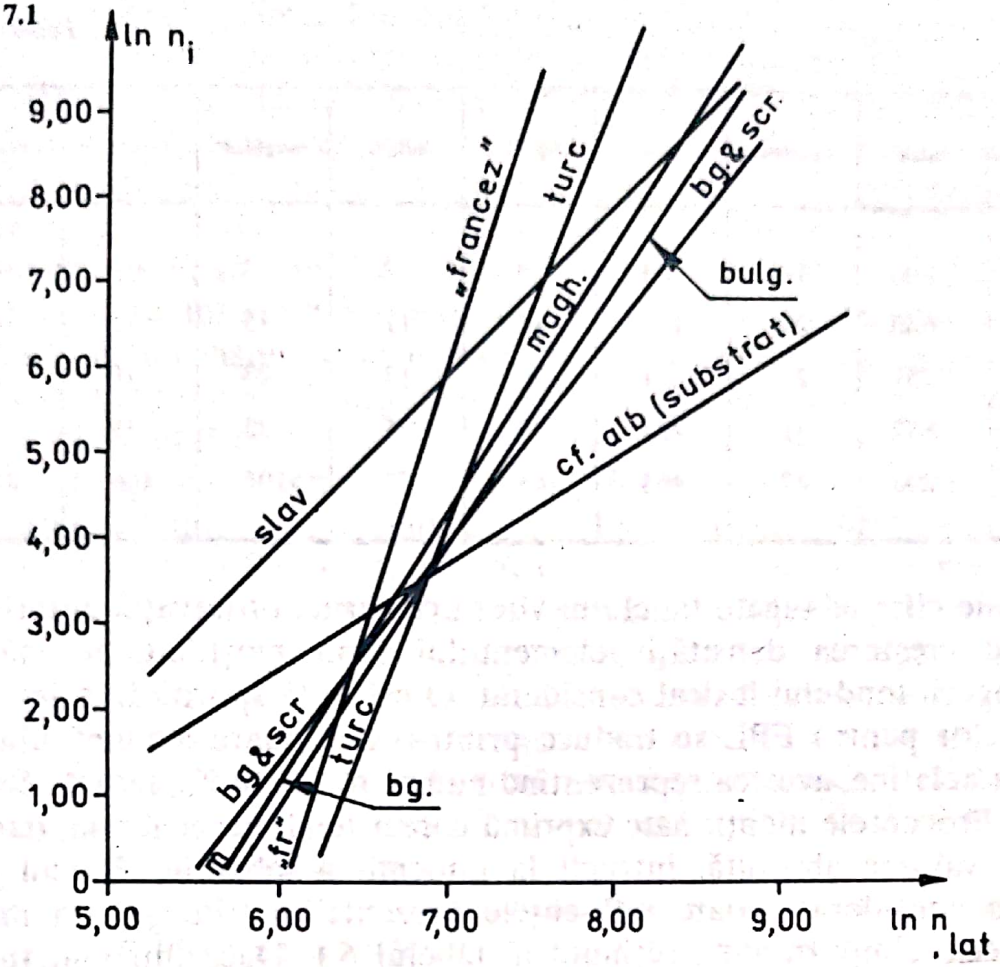
$$n_{tc} = exp\,(5{,}14\ ln\ n_{lat} - 31{,}75) \text{ pentru elementul turcesc}$$

$$nf_r = exp\,(6{,}36\ ln\ n_{lat} - 38{,}44) \text{ pentru elementul „francez"} \qquad (7.1)$$



Scanned with OKEN Scanner

FIG. 7.1




unde $n_{lat}$, $n_a$, $n_{sl}$, $n_{bg.scr.}$, $n_{bg}$, $n_m$, $n_{tc}$ şi $n_{fr}$ reprezintă numărul de cuvinte de origine latină şi, respectiv, comune cu albaneza, vechi slave, bulgaro-sîrbocroate, bulgăreşti, maghiare, turceşti şi „franceze".

Ecuaţiile (7.1) oferă un răspuns la întrebarea cîte unităţi lexicale de diferite provenienţe va conţine o colecţie de cuvinte de tip „vocabular fundamental" (v. definiţia în nota de la pag. 31), dacă ea cuprinde $n_{lat}$ termeni latini moşteniţi. Fie, de pildă, două asemenea mulţimi situate în afara domeniului acoperit de variantele FPL, una mai restrînsă $F_0$, cuprinzînd numai 400 de cuvinte de origine latină, şi alta mai extinsă $F_4$, conţinînd de trei ori mai multe. Cu ajutorul formulelor (7.1) am calculat numărul de unităţi de diferite origini pe care ar fi de aşteptat să le întîlnim în $F_0$ şi $F_4$. Rezultatele figurează în tabelul 7.1 alături de distribuţia după provenienţă[1] a cuvintelor din variantele de fond principal $F_1$, $F_2$, şi $F_3$.

[1] E vorba de valorile *calculate* (tot cu ecuaţiile (7.1)), iar nu de cele *inventariate* în tabelul 6.1. care diferă uşor de primele, în măsura în care modelul matematic diferă de realitatea modelată.



Scanned with OKEN Scanner

*Tabelul 7.1*

| | latin | substrat | slav | bg.scr. | bulgar | maghiar | turc | „francez" |
|---|---|---|---|---|---|---|---|---|
| $F_0$ | 400 | 12 | 51 | 4 | 2 | 3 | — | — |
| $F_1$ | 660 | 22 | 139 | 14 | 11 | 15 | 5 | 17 |
| $F_2$ | 751 | 26 | 181 | 20 | 17 | 23 | 10 | 39 |
| $F_3$ | 852 | 31 | 234 | 28 | 25 | 35 | 19 | 91 |
| $F_4$ | 1 200 | 47 | 464 | 68 | 72 | 104 | 109 | 775 |

Noile cifre adăugate tabelului vin să confirme observaţia anterioară privind creşterea densităţii elementului latin moştenit pe măsura restrîngerii fondului lexical considerat. O exigenţă sporită în selectarea cuvintelor pentru FPL se traduce printr-o diminuare a numărului de unităţi nelatine, acestea reprezentînd numai 16,1% în $F_0$, faţă de 58,7% în $F_4$. Procentele menţionate exprimă corect tendinţa generală, dar ele nu au valoare absolută, întrucît la întocmirea tabelului 7.1 nu s-au luat în considerare toate influenţele prezente în FPL şi nici măcar toate cele „importante", reţinute în tabelul 6.1. Dacă din punctul de vedere al scopului urmărit în acest paragraf, unde ceea ce interesează este numai poziţia *relativă* a diferitelor componente istorice ale vocabularului, simplificarea operată nu influenţează concluziile analizei, în schimb atunci cînd se va pune problema generală a dimensionării „vocabularelor fundamentale de extensiune dată" ne va fi interzis să facem abstracţie măcar şi de cel mai neînsemnat idiom implicat în formarea lexicului românesc.

Cum se prezintă, potrivit tabelului 7.1, ordinea de importanţă a surselor etimologice ale limbii române?

Examinînd situaţia din $F_0$, am putea afirma că, după elementul latin, determinant, ierarhia aporturilor lexicale străine este: 1. elementul vechi slav, 2. substratul, 3. elementul bulgar (inclusiv unităţile cu etimologie mixtă bulgaro-sîrbocroată), 4. elementul maghiar, în timp ce influenţa turcă şi cea neologică („franceză") sînt nedecelabile, dat fiind că termenii de originile respective nu au reuşit să pătrundă suficient de adînc în intimitatea vocabularului de bază.

Cu totul alta este situaţia dacă judecăm după imaginea oferită de $F_4$. Elementul „francez", inexistent în $F_0$, ocupă aici nici mai mult nici



Scanned with OKEN Scanner

mai puţin decît primul loc după fondul latin, în vreme ce substratul se vede azvîrlit pe ultima poziţie a clasamentului. Împrumuturile turceşti, care nu aveau nici un reprezentant în $F_0$, au ajuns să le depăşească pe cele maghiare, în ciuda ascensiunii rapide a acestora (numărul cuvintelor de origine maghiară a crescut de 10 ori mai repede decît al celor latine moştenite!). Dinamica deloc banală a acestor raporturi, pe lîngă faptul că oferă un generos subiect de meditaţie, ne permite şi să întrezărim o cale de conciliere a unor disensiuni deja seculare în lingvistica românească. Se ştie că inventarierea pe origini a cuvintelor româneşti iniţiată de A. de Cihac (constatarea sa, potrivit căreia elementul slav ar fi de două ori mai puternic decît cel latin, a provocat riposta binecunoscută a lui Hasdeu) şi reluată periodic pe corpusuri de întindere variabilă a condus de fiecare dată la rezultate diferite. Convinşi fiecare de adevărul lor, unii lexicologi au mers chiar pînă la a denigra, cu o nedreaptă pornire, strădania predecesorilor. Astfel, Al. Graur, absolutizînd importanţa „fondului principal" pe care îl socotea singurul în măsură să dea seamă de importanţa reală a diverselor influenţe, observa în legătură cu concluzia lui D. Macrea că „elementul latin din vocabularul românesc însumează în dicţionar 4,85 la sută": „Falsitatea acestui rezultat apare limpede dacă îl punem faţă în faţă cu statistica mea".

În realitate, deosebirile dintre rezultatele statisticilor n-au nimic de a face cu categoriile de adevăr şi fals şi nici cu probitatea ştiinţifică a celor ce le-au întocmit, calitate pe care nici nu se cuvine să o punem în discuţie cînd ne referim la nume ca A. de Cihac, Hasdeu, Sextil Puşcariu, D. Macrea sau Al. Graur. Toate statisticile citate sînt *adevărate* în măsura în care reflectă onest o situaţie reală şi nimic nu ne îndreptăţeşte să o socotim pe una dintre ele *mai adevărată* decît celelalte. Variaţia exponenţială a raporturilor dintre submulţimile de cuvinte de diferite provenienţe explică de ce repartiţia procentuală nu se menţine neschimbată dacă extindem sau restrîngem „cercul" vocabularului considerat. Componentele latină şi slavă, comparabile în bilanţul întocmit de D. Macrea după DLRM, stau în raportul de 4:1 la Al. Graur, ajungînd pentru $F_0$ la 8 : 1. Într-un vocabular fundamental conţinînd numai 200 sau 100 de cuvinte latine, n-am mai găsi, potrivit ecuaţiei (6.1), decît 14 şi, respectiv, 3 cuvinte de origine slavă, ceea ce revine la raporturi de 14 : 1 şi 33 : 1. Care dintre aceste rezultate ne oferă imaginea *reală* a poziţiei relative a celor două influenţe? Evident, toate şi nici unul. Toate — în măsura în care reprezintă adevăruri „de nivel", perfect valabile în condiţiile în care au fost stabilite; nici unul — întrucît luat în parte, fiecare dintre aceste



Scanned with OKEN Scanner

rezultate se vădeşte insuficient pentru caracterizarea globală a relaţiei respective.

Raportul numeric dintre unităţile de diferite origini variază continuu cu extensiunea vocabularului considerat. Dar aceasta nu e totul. Exemplul de mai sus, privitor la relaţiile dintre elementul latin şi cel slav, ar putea crea falsa impresie că, între două idiomuri ce au participat la formarea lexicului românesc, există întotdeauna unul care îl domină pe celălalt, variabilă cu volumul vocabularului fiind numai proporţia acestei dominări. Lucrurile nu stau însă deloc astfel.

Pentru a facilita înţelegerea acestui aspect, să presupunem din nou acceptată ipoteza „liniarităţii"[1]. În baza ei, numărul de cuvinte de orice origine poate fi reprezentat în scară logaritmică în funcţie de numărul cuvintelor de obîrşie latină, sub forma unei drepte. Două componente etimologice distincte sînt caracterizate, prin urmare, de o pereche de drepte. Acestea nu pot fi, prin definiţie, decît fie concurente, fie paralele, fie suprapuse. Ultimele două situaţii implică pante egale $m_1 = m_2$, ceea ce nu am întîlnit la nici un cuplu de influenţe cercetate aici[2]. Reprezentate pe un acelaşi grafic, toate dreptele din figurile 6.1 — 6.8 sînt concurente (fig. 7.1). Punctul lor de intersecţie le împarte în semidrepte cu proprietatea că, la stînga acestui punct, dreapta cu pantă mai mică este situată deasupra celei cu pantă mai mare, în timp ce la dreapta lui cotele liniei mai abrupte le depăşesc pe ale celeilalte. Aşadar, dacă un element este numeric preponderent pînă în secţiunea intersecţiei, el va fi, în mod necesar, dominat începînd cu această secţiune.

Cunoscînd expresiile analitice ale dreptelor caracteristice, determinarea punctelor de intersecţie este extrem de simplă. Pentru relaţia latin *versus* slav cele două drepte au expresiile:

$$y = x \text{ şi } y = 2{,}01\,x - 8{,}11$$

Prin egalarea membrilor drepţi se obţine imediat $x \approx 8{,}03$, de unde $n_{lat} \approx exp\ 8{,}03 \approx 3\,080$, care este abscisa (egală cu ordonata, întrucît se află pe bisectoarea cadranului întîi) punctului de intersecţie. Interpretarea acestui rezultat este următoarea: un vocabular de tip „fond principal" cuprinzînd 3 080 de cuvinte latine, conţine totodată

---

[1] Ipoteza invocată nu constituie o condiţie indispensabilă demonstraţiei, dar posibilitatea reprezentării sub formă de drepte a raporturilor dintre influenţe simplifică raţionamentul, făcînd mai intuitiv obiectul discuţiei.

[2] Pantele dreptelor figurează în ecuaţii sub forma coeficienţilor argumentului $x$, care s-au dovedit diferiţi pentru toate relaţiile stabilite anterior.



Scanned with OKEN Scanner

3 080 de cuvinte slave; în orice vocabular mai restrîns decît acesta elementul latin rămîne precumpănitor, după cum în orice vocabular mai extins predomină elementul slav.

Procedînd la fel pentru raporturile latin *versus* turc, obţinem punctul de intersecţie $n_{lat} = n_{tc} = 2\,040$. Se regăseşte astfel rezultatul „absurd" al lui A. de Cihac: egalitatea ponderii numerice a fondului latin cu împrumuturile turceşti. Singura deosebire constă în pragul mai coborît la care survenea această egalitate în condiţiile vocabularului mai sărac cuprins în dicţionarul lingvistului ardelean: $n_{lat} = n_{tc} = 1\,150$. La drept vorbind însă, nu atît valoarea pragului contează (el nici nu putea fi acelaşi pentru etape atît de diferite ale istoriei vocabularului românesc cum sînt cele reflectate de lexiconul lui A. de Cihac şi de DLRM), cît existenţa în sine a unui asemenea prag, care explică situaţii ce li s-au părut înaintaşilor incomprehensibile sau de neacceptat.

Este ceea ce i s-a întîmplat lui Al. Graur, care, surprins de numărul cu adevărat impresionant de cuvinte de provenienţă „franceză" din *Dicţionarul limbii române din trecut şi de astăzi,* nu întrevedea altă ieşire din impas decît punerea la îndoială a corectitudinii etimologiilor propuse de A. I. Candrea: „N-am făcut socoteala (! n.n) dar sînt convins că elementele sigur franceze din limba noastră nu trec de 3 sau 4 la sută"[1].

Procentele, avansate fără argumente, sînt cam cele corespunzătoare ponderii elementului „francez" în FPL, ceea ce contrazice însăşi observaţia lui Al. Graur conform căreia „cînd mărim numărul total al cuvintelor, scade procentul elementelor latine şi creşte procentul celorlalte cuvinte". Dacă, aplicînd consecvent principiile de admitere în fondul principal pe care şi le-a impus, ar fi extins limitele acestuia pînă la includerea a cca 1 300 de cuvinte de origine latină, Al. Graur ar fi putut avea surpriza de a întîlni în mulţimea astfel alcătuită tot 1 300 de cuvinte de provenienţă „franceză", pentru ca mai departe fondul latin moştenit să rămînă în minoritate, întocmai ca în incriminatul „dicţionar cosmopolit al lui Candrea". Speculîndu-şi pînă la capăt intuiţiile, indiscutabil juste, privind ierarhia de importanţă a cuvintelor româneşti, Al. Graur putea găsi el însuşi explicaţia unui atare fenomen, dacă, din păcate, nu ar fi fetişizat conceptul de fond principal lexical, socotind că există o unică selecţie cu proprietatea de a fi „cea mai justă" şi „demnă de încredere cu privire la stabilirea

[1] *Încercare* ... pag. 36



Scanned with OKEN Scanner

procentelor pe limbile de origine"[1]. În viziunea sa, numai lista $F_2$ îndeplineşte această condiţie maximală.

Că $F_2$ constituie într-adevăr o variantă de vocabular fundamental stabilită judicios ne-o demonstrează observaţia că, în reprezentările noastre grafice (fig. 6.1—6.8), $F_2$ face parte întotdeauna din tripleta punctelor coliniare, atunci cînd o asemenea tripletă există, ceea ce nu se întîmplă în mod sistematic nici cu $F_1$ nici cu $F_3$ şi nici chiar cu D. Afirmaţia lui Al. Graur trebuie înţeleasă atunci nu în sensul că raporturile dintre influenţe sînt cele date de $F_2$ şi numai acestea, ci că ele variază după o curbă care, cu o mare probabilitate, trece prin punctul $F_2$.

## 8. Indicatori ai importanţei cuvintelor. Productivitatea lexicală. Corelaţia cu vechimea

După ce am examinat distribuţia pe origini a unităţilor cuprinse în variantele de FPL propuse de Al. Graur, extrăgînd din reprezentarea grafică a valorilor obţinute o posibilă lege cantitativă, să nu uităm a sublinia fragilitatea acestui rezultat. Ispititoare prin consecinţele pe care le-ar putea avea asupra întregii problematici a organizării vocabularului, „ipoteza liniarităţii", cum vom continua să o numim, nu înseamnă, pentru moment, nimic mai mult decît o simplă ipoteză. Aceasta din pricină că, pe de o parte, s-a bazat pe un material faptic prea sărac (numai patru secţiuni prin piramida lexicală, corespunzătoare colecţiilor $F_1$, $F_2$, $F_3$ şi mulţimii D a tuturor cuvintelor din DLRM) şi că, pe de altă parte, însăşi maniera de obţinere a acestor secţiuni era copios marcată de subiectivism, trei dintre ele constituind opţiuni personale ale unui lingvist anume (Al. Graur). Încă insuficient fundamentat pentru a putea revendica statutul de lege lingvistică, acest rezultat empiric trebuie privit deocamdată doar ca o sugestie privind direcţia în care ar fi de căutat un eventual principiu de structurare a lexicului românesc.

Înainte de a păşi mai departe, să ne reamintim criteriile de admitere a cuvintelor în FPL, aşa cum au fost ele enunţate de Al. Graur.

[1] *Op. cit.* pag. 62.



Scanned with OKEN Scanner

Exceptînd condiţia, oricum subînţeleasă, a uzului general, ele ar fi în număr de 4 şi anume:

a. Productivitatea lexicală
b. Polisemia
c. Participarea la expresii şi locuţiuni
d. Frecvenţa de folosire

Aplicat separat, fiecare criteriu conduce la un „clasament" al cuvintelor în general diferit de cel obţinut pe baza oricăruia dintre celelalte trei criterii. Se obţin astfel patru liste distincte de cuvinte, problema identificării fondului principal revenind la contopirea lor într- o listă unică şi la stabilirea unei frontiere care să delimiteze FPL de masa vocabularului.

Enunţată în acest fel, problema pare simplă, dar rezolvarea ei se izbeşte de dificultăţi imense. Riguros vorbind, în măsura în care se prezintă ca o problemă de *agregare a ierarhiilor*[1] ea este chiar insolubilă (cel puţin pentru condiţiile de „raţionalitate" formulate de Arrow în celebra sa teoremă). Nu insistăm asupra acestui aspect, la care vom reveni la momentul potrivit. Menţionăm numai existenţa unor obstacole *de principiu* în calea ordonării după importanţă a vocabularului, operaţiune care presupune sinteza mai multor clasamente diferite, ceea ce nu se poate realiza fără imixtiunea unor elemente de decizie subiective („dictatoriale" în terminologia domeniului[2]). Înainte de a aborda această latură a problemei, trebuie să parcurgem, cum este şi firesc, etapa pregătitoare, indispensabilă, a stabilirii celor patru clasamente ale cuvintelor în conformitate cu criteriile enunţate. Examinînd modul cum se structurează lexicul românesc în funcţie de fiecare criteriu luat în parte, vom putea analiza atunci compatibilitatea acestor criterii cu „ipoteza liniarităţii" şi depista pe această cale cauza apariţiei unei dependenţe de tip liniar (recte exponenţial) în variantele de fond principal propuse de Al. Graur.

Vom considera aşadar succesiv cîte un singur criteriu şi pe baza lui vom ordona cuvintele ca şi cum importanţa lor nu ar depinde decît de acest unic parametru, urmînd să tragem toate concluziile ce decurg dintr-o atare ipoteză.

---

[1] Referitor la această chestiune cf. Gh. Păun: *Paradoxurile clasamentelor,* Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1987.

[2] O decizie „dictatorială", în sensul definiţiei matematice, este şi selecţia lui Al. Graur, de care în acest moment ne despărţim.



Scanned with OKEN Scanner

Începem cu productivitatea lexicală, parametru mai lesne cuantificabil, întrucît derivatele unui cuvînt sînt mai uşor de identificat şi de numărat decît, bunăoară, sensurile lui. Lectura repetată şi atentă a lucrării lui Al. Graur ne-a convins de altfel că şi în viziunea autorului bogăţia de derivate constituie un criteriu privilegiat în raport cu celelalte. Să mai adăugăm la aceasta şi însemnătatea în plan diacronic a productivităţii, ştiut fiind faptul că între vechimea cuvintelor şi numărul derivatelor lor subzistă o strînsă dependenţă.

Pentru a obţine o imagine cît mai cuprinzătoare a corelaţiei provenienţă-productivitate, am grupat pe origini şi familii lexicale toate cele aproape 50 000 de cuvinte din DLRM. Din păcate, ceea ce părea iniţial un avantaj şi anume faptul că dispuneam ca mijloc de verificare de statistica datorată lui D. Macrea s-a dovedit în urmă a fi nu un sprijin, ci mai degrabă o sursă de nedumerire şi îngrijorare, deoarece între inventarul nostru şi cel al autorului menţionat apar deosebiri importante.

Tabelul 8.1 reţine cifrele globale privind unităţile primare (cuvinte-bază) şi cele de al doilea rang (derivate şi variante) aparţinînd principalelor 20 de componente istorice ale vocabularului românesc, selectate dintre cele 118 prezente în statistica noastră[1]. Nepotrivirile dintre cele două serii de cifre nu privesc atît numărul total de cuvinte de diferite origini, cît, mai ales, proporţia dintre cuvintele-bază, pe de o parte, şi derivatele şi variantele acestora, pe de alta. În mod sistematic, noi am găsit mai multe (pentru situaţia elementului „francez" cu aproape 2 300!) cuvinte-bază, dar şi mai puţine derivate (în acelaşi caz cu peste 2 500!). Nu ne putem explica anomalia decît prin înregistrarea la D. Macrea ca derivate formate pe teren românesc a unor împrumuturi aparţinînd în limba sursă unei aceleiaşi familii lexicale. Astfel cuvintele *colonial, colonialism, colonialist, coloniza, colonizator, coloniale, colonist* nu provin în româneşte din neologismul *colonie,* ci au fost împrumutate independent, unele chiar din limbi diferite (*colonist* ⟨ germ. *kolonist*). Includerea lor abuzivă în familia lui *colonie* induce în eroare, conferindu-i acestui cuvînt o aparenţă de prolificitate, cînd, în realitate, progenitura sa se limitează la un singur

[1] La D. Macrea apar numai 76 de grupe, diferenţa explicîndu-se şi prin înregistrarea separată de către noi a elementelor nesigure la rubrici diferite, în funcţie de proximitatea etimologică indicată de dicţionar: astfel, de pildă, cuvintele *obcină, sliboviţă, tască* figurează la noi în categoriile distincte cf. *ceh, rus,* cf. *scr, pol.* şi respectiv, cf. *magh, ucr,* în timp ce D. Macrea nu a prevăzut pentru înmatricularea lor decît rubrica unică „origine incertă".



Scanned with OKEN Scanner

descendent: *semicolonie*. Aşa se face că, dacă în ceea ce priveşte totalul aportului „francez", diferenţa dintre inventarul nostru şi cel al lui D. Macrea abia dacă depăşeşte un procent, raportul dintre derivate + variante şi cuvintele-bază diferă, în schimb, de la simplu la dublu (0,25 faţă de 0,50). Or, citarea valorilor acestui raport în versiunea D. Macrea a devenit curentă în lucrările româneşti de istorie a limbii şi faptul că cifrele respective nu corespund realităţii nu poate trece nesemnalat.

*Tabelul 8.1.*

| | | Latine | Slave | Bulgare | Bulg. & Scr. | Turc. | Magh. | Neogr. | Franc. | Lat. Sav. |
|---|---|---|---|---|---|---|---|---|---|---|
| M. Dinu | Cuvinte-bază | 1 849 | 1 133 | 279 | 212 | 1 048 | 443 | 579 | 12 770 | 738 |
| | Derivate şi variante | 8 071 | 2 826 | 602 | 535 | 779 | 630 | 593 | 6 359 | 448 |
| | TOTAL | 9 920 | 3 959 | 881 | 747 | 1 827 | 1 073 | 1 172 | 19 129 | 1 186 |
| D. Macrea. | Cuvinte-bază | 1 794 | 1 038 | 292 | 200 | 1 089 | 448 | 615 | 15 051 | 800 |
| | Derivate şi variante | 7 642 | 2 877 | 583 | 549 | 759 | 587 | 590 | 3 819 | 451 |
| | TOTAL | 9 436 | 3 915 | 875 | 749 | 1 848 | 1 037 | 1 205 | 18 870 | 1 251 |

| | | Ital. | Germ. | Onom. | Necun. | Cf alb.. | Rus. | Scr. | Ucr. | Engl. | Săseşti | Ţigăn. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M. Dinu | Cuvinte-bază | 430 | 561 | 272 | 1440 | 81 | 252 | 203 | 193 | 60 | 18 | 17 |
| | Derivate şi variante | 421 | 317 | 836 | 1330 | 380 | 200 | 238 | 205 | 13 | 19 | 7 |
| | TOTAL | 851 | 878 | 1108 | 2770 | 461 | 452 | 441 | 398 | 73 | 37 | 24 |
| D. Macrea | Cuvinte-bază | 480 | 619 | 281 | 1764 | 83 | 324 | 214 | 207 | 67 | 18 | 14 |
| | Derivate şi variante | 238 | 292 | 766 | 1122 | 351 | 182 | 233 | 181 | 10 | 14 | 9 |
| | TOTAL | 718 | 911 | 1047 | 2886 | 434 | 506 | 447 | 388 | 77 | 32 | 23 |



Scanned with OKEN Scanner

Fireşte, nu ne arogăm privilegiul infailibilităţii. Este foarte posibil, ba chiar probabil, avînd în vedere volumul mare de lucru, ca şi în statistica noastră să existe scăpări, dar faptul că numărul total de cuvinte găsit de noi este aproape identic cu cel stabilit de D. Macrea arată că erorile, cîte există, nu sînt cu precădere de natura omisiunilor. Deosebirile par să provină dintr-o tendinţă a lui D. Macrea de a nu se limita la atribuirile din dicţionare, cu care bănuim că n-a fost întotdeauna de acord, şi din încercarea de a le corecta, ceea ce l-a condus în final la o statistică interpretativă în locul unui inventar impersonal, obiectiv. În favoarea acestei teze pledează numărul mai mic de unităţi de origine necunoscută consemnate de D. Macrea, semn că lingvistul a tranşat într-un fel propriu o parte dintre controversele etimologice care i-au determinat pe autorii dicţionarului (revizuit, totuşi, de însuşi D. Macrea!) să se abţină de la un diagnostic univoc. Tot unei opţiuni personale pare să i se datoreze şi absenţa din statistica lui D. Macrea a unei clase a cuvintelor compuse, deşi acestea constituie o categorie deloc neglijabilă. Potrivit numărătorii noastre, 453 de cuvinte cu 119 derivate şi 70 de variante, adică un total de 642 de unităţi din dicţionar au luat naştere prin compunere. Dintre acestea, un număr de 397 (250 de cuvinte-bază, 86 de derivate şi 61 de variante) sînt alcătuite exclusiv din elemente latine moştenite. După exemplul lui Al. Graur, noi le-am rezervat o rubrică aparte, inexistentă la D. Macrea, care, bănuim, avînd în vedere numărul mai mare de cuvinte de origine latină ce i-au rezultat, că le-a asimilat acestora. Restul, constînd din 203 cuvinte-bază cu 33 de derivate şi 9 variante este alcătuit din elemente de cele mai variate provenienţe şi este greu de spus unde trebuie căutate ele în statistica lui D. Macrea. În cazul că le-ar fi inclus în categoria cuvintelor cu dublă origine, inventarul său ar fi trebuit să cuprindă, de exemplu, şi tipuri ca *latin + ucrainian (ghemotoc = ghem + motoc (el))* sau *cf. albanez + sas (ciocofleandură = cioc + fleandură)*, ceea ce însă nu se întîmplă. Elucidarea completă a modului cum a procedat autorul statisticii este aşadar imposibilă, dacă ne limităm, şi n-avem cum face altminteri, exclusiv la informaţiile cuprinse în deştul de laconica prezentare din lucrarea citată.

Oricum, demersul lui D. Macrea pare să fi constat numai în simpla numărare a cuvintelor, în timp ce noi le-am extras din dicţionar şi grupat pe familii, operaţiuni ce diminuează riscul de eroare şi, în orice caz, fac mai improbabilă omiterea unor cuvinte productive, căci scăparea din vedere a acestora are toate şansele de a fi corectată de



Scanned with OKEN Scanner

îndată ce se întîlneşte primul derivat al unităţii pierdute. Or, cum ceea ce ne interesează aici sînt tocmai familiile lexicale, analiza noastră nu se află în pericol de a fi denaturată de omisiuni majore. Cu cît este mai productiv, cu atît e mai dificil pentru un cuvînt să scape prin ochiurile statisticii, întrucît fiecare derivat al său ni-l reaminteşte şi nu este posibil să-l rătăcim decît laolaltă cu întreaga familie pe care a generat-o, ceea ce pare mai puţin probabil.

În tabelul 8.1 am adăugat derivatelor variantele de pronunţare menţionate în dicţionar, din nevoia de a asigura comparabilitatea rezultatelor cu cele ale lui D. Macrea, care nu a operat distincţia dintre cele două categorii de cuvinte. În continuare însă vom separa aceste două clase, interesîndu-ne numai de derivate, singurele importante deopotrivă pentru caracterizarea locului în sincronie al cuvintelor matcă din care provin şi pentru stabilirea vechimii lor în limbă. Din ambele puncte de vedere numărul de variante nu are o semnificaţie comparabilă, căci nu reflectă nici însemnătatea cuvîntului (adesea tocmai frecvenţele şi ariile mai restrînse de folosire favorizează menţinerea unor rostiri diferite) şi nici vîrsta sa (sînt binecunoscute oscilaţiile de pronunţare a neologismelor, mai cu seamă a celor provenite din limbi cu sisteme fonologice mult diferite de al nostru, existenţa variantelor fonetice fiind uneori chiar un indiciu că împrumutul este foarte recent, încă neasimilat).

După cum am mai spus-o, cercetarea noastră s-a axat pe materialul cuprins în DLRM, fiindcă ne-am propus caracterizarea lexicului românesc de la mijlocul secolului al XX-lea, moment cînd apare proiectul de vocabular fundamental al lui D. Graur. Aceasta nu însemna însă că trebuia să ne mulţumim cu stadiul de cunoaştere lingvistică oglindit de dicţionarul respectiv. În cele trei decenii care au urmat apariţiei DLRM, studiul etimologiilor a înregistrat unele progrese, astfel încît o parte din incertitudinile autorilor lucrării din 1958 s-au disipat. Ni s-a părut natural să ţinem seamă de aceste clarificări, în care scop pentru toate situaţiile lăsate nedecise în DLRM am apelat la echivalentul mai recent al acestei lucrări care este Dicţionarul Explicativ al Limbii Române (Editura Academiei 1984). Ca urmare, rubricile rezervate diferitelor provenienţe s-au mai îmbogăţit, pe seama restrîngerii clasei elementelor cu etimologie necunoscută.

Tot cu ajutorul DEX am tranşat şi o altă categorie de probleme de atribuire, nu mai puţin spinoasă, şi anume cea care priveşte

Scanned with OKEN Scanner

componenta slavă a vocabularului. Este şi acesta un domeniu în care mai stăruie multe incertitudini. Împărţind, aşa cum a făcut-o şi D. Macrea, numărul de derivate şi variante la numărul de cuvinte-bază, pentru fiecare dintre constituenţii etimologici ai lexicului românesc, se obţin valorile din tabelul 8.2.

*Tabelul 8.2.*

| | Lat. | Slav. | Bulg. | Bg. şi scr. | Turc | Magh. | Ngr. | „Franc" | Lat. sav. | Ital. |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv-bază: Nr. derivate şi variante | 4,26 | 2,77 | 2,00 | 2,75 | 0,70 | 1,31 | 0,96 | 0,25 | 0,56 | 0,50 |

| | Germ. | Onom. | Nec. | Cf. alb. | Rus | Scr. | Ucr. | Engl. | Sas. | Ţig. |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv-bază: Nr. derivate şi variante | 0,47 | 2,73 | 0,64 | 4,23 | 0,56 | 1,09 | 0,87 | 0,15 | 0,78 | 0,64 |

Observaţia potrivit căreia cuvintele mai vechi au generat în medie mai multe derivate şi variante decît împrumuturile mai recente se confirmă şi pentru bilanţul întocmit de noi. Între fondul lexical moştenit (cuvintele latine şi cele de substrat sînt aproape egal de productive) căruia îi corespund valorile maxime ale raportului respectiv şi componentele neologice „franceză" şi engleză ale vocabularului, caracterizate de valorile cele mai scăzute ale acestuia (0,25 şi, respectiv, 0,15), cele 20 de influenţe luate în considerare se înşiră descrescător într-o ordine foarte apropiată de cea cronologică a exercitării lor. Vom avea mai tîrziu prilejul de a analiza pe larg această relaţie. Deocamdată nu vrem decît să evidenţiem faptul că împrumuturile cărora dicţionarul le dă o dublă explicaţie etimologică prin bulgară şi sîrbocroată se comportă sub aspectul generării de derivate şi variante practic identic cu elementele vechi slave (un raport de 2,75 faţă de 2,77), ceea ce ar sugera că şi în primul caz avem de a face cu împrumuturi mai vechi, probabil anterioare separării din trunchiul unic sud-slav a grupului limbilor de est (bulgara şi macedoneana) de al celor de vest



Scanned with OKEN Scanner

(sîrbocroata şi slovena)[1]. În sprijinul acestei teze vine şi observaţia lui Dorin Gămulescu[2] potrivit căreia dacă examinăm cu atenţie aceste elemente vom constata că: „Marea majoritate sînt vechi sud-slave *(baştină, buruiană, cetină, chişiţă, cîrpă, claie, cobiliţă, cocean)* şi numai o parte au etimologie dublă, ca de pildă *cocină*" (op. cit. pag. 55). Dicţionarul Explicativ merge şi el parţial pe această linie, modificînd în *slav* indicaţia etimologică *bulgar, sîrb* din DLRM pentru un număr de 13 cuvinte, aproape în totalitate productive (cifrele din paranteză reprezintă numărul de derivate ale fiecăruia): *colac* (9), *năuc* (6), *năvăli* (7), *ocărî* (3), *pilă* (5), *pintenog, piron* (3), *plată* (2), *podmet* (reg), *poiată* (1), *ponor* (3), *sinod, ţigan* (17). Dificultatea de a distinge clar între cele două valuri de împrumuturi este reflectată şi de etimologia de compromis propusă de DEX pentru cuvinte ca *proţap* (4) < sl. *pročepu*, bg. *procep, pîrtie* < sl. *prŭtŭ* scr. *prt, pisar* < sl. *pisari*, pol. *pisarz, polog* < sl. *podlogu* rus. *polog, ştevie* < sl. *štavije*, bg. *štavel*, scr. *štavlije*, şi chiar *podnojie* < sl. *podnožije*, bg. *podnega* scr. *podnožje*, rus *podnožje*, ucr. *podžžja*. Că enumerarea acestora nu închide lista unităţilor care trebuie considerate de fapt vechi slave ne-o dovedeşte situaţia unor cuvinte comune cu aromâna cum ar fi *cocean* (bg., scr. atît în DLRM cît şi în DEX), *cosiţă* (2) (bg. în ambele dicţionare), *grădină* (8) (bg., scr. în DLRM şi în DEX), *livadă* (bg., ucr. în ambele dicţionare), *padină* (1) (bg., scr. în DLRM şi bg. în DEX), *opări* (5) (bg., scr. în ambele dicţionare). Bazat pe faptul, istoriceşte demonstrat, că ruperea contactului dintre românii din Peninsula Balcanică şi cei locuind de-a lungul şi la nordul Dunării a avut loc pe la sfîrşitul secolului al X-lea, A. Rosetti socoteşte, pe bună dreptate, că aceste cuvinte aparţin aportului slav anterior cristalizării limbilor bulgară şi sîrbocroată ca entităţi lingvistice de sine stătătoare[3]. În orice caz, operarea unei disctincţii clare între aportul vechi slav şi contribuţia mai tîrzie a limbilor slave meridionale rămîne dificilă. D. Gămulescu[4] o spune chiar mai tranşant: „Este, deci, imposibil ca, sprijinindu-ne în exclusivitate pe materialul lingvistic, să stabilim o linie netă de demarcaţie între perioada veche sud-slavă şi cea a începutului influenţei moderne".

---

[1] Ivan Popovic (*Geschichte der serbokroatischen Sprache,* Wiesbaden, Otto Harrasowitz, 1960, pag. 232) consideră că unitatea s-a menţinut pînă pe la anul 1000.

[2] *Contribuţii la studierea elementelor de origine sîrbocroată ale vocabularului daco-român,* Bucureşti — Pančevo, 1974, p. 55.

[3] A. Rosetti *Istoria limbii române* III, Ed. Ştiinţifică, 1964, p. 58—59.

[4] *Op. cit.* p. 55.



Scanned with OKEN Scanner

Conştienţi că nu noi sîntem cei chemaţi să aducem lumină în acest domeniu, am dorit ca datele cu care operăm să reflecte stadiul cel mai recent de cunoaştere a problemei şi de aceea ne-am adresat din nou Dicţionarului Explicativ în calitate de cea mai nouă lucrare lexicografică academică încheiată ce conţine şi indicaţii etimologice. Reexaminarea în lumina DEX a etimologiilor propuse în DLRM pentru întregul domeniu lexical neo-slav a relevat nu mai puţin de 339 de deosebiri. Un adevărat carusel al atribuirilor, în care, în afara identificării unor etimonuri vechi slave pentru cuvinte socotite în DLRM drept achiziţii lexicale mai recente, au loc nenumărate rocade: *juli* şi *odolean* trec din domeniul sîrbocroat în cel bulgar şi, respectiv, ucrainian, elementul rus pierde pe *marchitan* în favoarea celui polonez şi pe *năclăi* în avantajul componentei bulgare a lexicului, care mai cîştigă şi pe ex-ucrainianul *pripor* şi aşa mai departe. Unele transferuri ne scot chiar în afara spaţiului slav: *mierţă, no* şi *ogar* trec de la sîrbocroate la maghiare, *mandră* şi *spermanţet* de la bulgare şi, respectiv, ruseşti, la împrumuturile din neogreacă, *îngurlui* şi *picni* sînt acum recunoscute drept onomatopeice, după ce au fost considerate unul bulgăresc, altul sîrbocroat etc. Mai curios, la unele etimologii avansate fără dubiu (nu sub protecţia acoperitorului *cf*!) se renunţă cu totul în DEX, astfel încît termenii respectivi vin să îngroaşe rîndurile cuvintelor de origine necunoscută. Aşa se întîmplă cu *măcriş, miezuină, pîrdalnic, tîrsoacă* (bulgare în DLRM), *ban* (monedă), *bulvan, mîndalac, mogîldeaţă, odgon* (sîrbocroate în aceeaşi lucrare), *maldăr* (polonez).

Bilanţul final rezultat în urma remanierilor efectuate după redistribuirea în conformitate cu indicaţiile din DEX a cuvintelor care figurau în DLRM ca avînd etimologie necunoscută şi a celor aparţinînd domeniului lexical slav este prezentat sintetic în tabelul 8.3. Şi de această dată au fost reţinute numai principalele 20 de influenţe (care acoperă însă peste 97% din numărul total al cuvintelor consemnate în DLRM).

Cifrele par să confirme existenţa unei legături directe între productivitate şi vechime. Valorile cele mai ridicate ale raportului $\delta$ dintre numărul de derivate şi cel de cuvinte-bază corespund substratului (cuvinte comune cu albaneza) şi fondului latin moştenit. Ele nu diferă între ele decît cu cca 1%, dar trebuie observat că ordinea valorică contravine ordinii cronologice, în sensul că elementul autohton se comportă ca şi cum l-ar fi precedat *în vocabularul românesc* pe cel latin, ceea ce nu concordă cu realitatea istorică. În această privinţă chiar termenul de substrat poate genera confuzie, în măsura în care



Scanned with OKEN Scanner

*Tabelul 8.3.*

| | Latine | Slave | Bulgare | Bg. şi Scr. | Turc. | Magh. | Neogr. | „Franc." | Lat. sav. |
|---|---|---|---|---|---|---|---|---|---|
| Cuvinte-bază | 1 802 | 1 101 | 287 | 172 | 1 092 | 453 | 624 | 15 088 | 801 |
| Derivate | 7 030 | 2 668 | 536 | 386 | 534 | 491 | 457 | 3 162 | 370 |
| Total | 8 832 | 3 769 | 823 | 558 | 1 626 | 944 | 1081 | 18 250 | 1 171 |
| $\delta = \frac{\text{derivate}}{\text{cuv.-bază}}$ | 3,900 | 2,423 | 1,868 | 2,244 | 0,489 | 1,084 | 0,732 | 0,210 | 0,462 |

| | Ital. | Germ. | Onom. | Necun. | cf. Alb. | Rus. | Scr. | Uer. | Engl. | Sas. | Ţigan. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Cuvinte-bază | 485 | 636 | 343 | 1 508 | 85 | 274 | 167 | 207 | 69 | 19 | 14 |
| Derivate | 142 | 241 | 663 | 856 | 336 | 109 | 125 | 145 | 4 | 9 | 8 |
| Total | 627 | 877 | 1 006 | 2 364 | 421 | 383 | 292 | 352 | 73 | 28 | 22 |
| $\delta = \frac{\text{derivate}}{\text{cuv.-bază}}$ | 0,293 | 0,379 | 1,933 | 0,568 | 3,952 | 0,398 | 0,749 | 0,700 | 0,058 | 0,474 | 0,571 |

sugerează preexistenţa *în limbă* a cuvintelor în cauză, cînd de fapt ele au preexistat numai *pe teritoriul* ei de formare. În raport cu limba română aflată în secolele II—III în stadiul de latină dunăreană, termenii zişi „de substrat" constituie însă de fapt împrumuturi, elemente noi de vocabular (neologisme!) pătrunse din limba băştinaşilor ca urmare a contactului şi apoi a fuziunii celor două populaţii. Or, dacă în medie cuvintele sînt cu atît mai productive cu cît vechimea lor este mai mare, ne-am fi aşteptat ca parametrul $\delta$ să înregistreze pentru elementul autohton o valoare mai mică decît cea corespunzătoare fondului lexical latin. Lucrurile nu stau aşa, dar acest fapt nu poate fi socotit drept o infirmare a corelaţiei prezumate, din cel puţin două motive:

1. Cuvintele comune cu albaneza nu sînt singurii termeni de substrat conservaţi de limba română (I. I. Rusu[1], de pildă, enumeră cca. 90 de unităţi lexicale „sigure sau probabil autohtone (preromane) care lipsesc în albaneză"). Luarea în considerare şi a acestora sau numai a cîtorva dintre ele modifică valoarea medie a productivităţii cuvintelor de substrat, putînd-o coborî sub cea aferentă elementului

[1] *Limba traco-dacilor*, ediţia a II-a revăzută şi adăugită, Ed. Ştiinţifică, Bucureşti, 1967, pag. 215—216.



Scanned with OKEN Scanner

latin moştenit. Este, bunăoară, suficient să adăugăm numai doi termeni neproductivi la cele 85 de cuvinte comune cu albaneza extrase din dicţionar, pentru ca raportul $\delta$ să scadă de la 3,952 la 3,862, adică mai puţin decît productivitatea medie a cuvintelor de origine latină. În condiţiile cunoaşterii încă insuficiente a aportului lexical traco-dacic la formarea vocabularului românesc, este aşadar riscantă orice absolutizare a cifrelor stabilite pe baza informaţiilor oferite de dicţionare, cu atît mai mult cu cît acestea nu constituie lucrări de cercetare etimologică propriu-zisă, ci urmăresc alte scopuri, indicaţiile privind originea cuvintelor reprezentînd în cazul lor doar un element de însemnătate secundară, tratat la un nivel de exigenţă corespunzător.

2. Vechimea medie a cuvintelor de substrat ar putea fi eventual obiectiv mai mare decît cea a fondului lexical latin, dacă se ţine seama că asimilarea acestor elemente trebuie să fi avut loc în prima fază a procesului de romanizare, de-a lungul perioadei relativ scurte de bilingvism daco-roman. Latina vorbită de băştinaşi în acest interval cînd ei deprindeau treptat limba ocupantului conţinea fără doar şi poate mai multe „dacisme" decît cele ce ni s-au păstrat. Într-o fază ulterioară, de durată probabil mult mai mare decît cea dintîi, presiunea continuă a limbii oficiale şi dispariţia ca mijloc curent de comunicare a idiomului indigen a făcut ca, treptat, relictele lexicale daco-getice să fie înlocuite cu sinonime latine, unităţile de substrat care au reuşit să se menţină în limbă apărînd astfel drept mai vechi decît cel puţin o parte dintre cuvintele latine moştenite. Scenariul se poate, desigur, complica prin luarea în considerare a unei posibile influenţe lingvistice, ulterioare retragerii aureliene, exercitate de limba vorbită încă în spaţiul de vieţuire a dacilor liberi. Prin pătrunderea acestora pe teritoriul fostei provincii, se va fi reîmprospătat poate întrucîtva zestrea lexicală autohtonă a latinei dunărene. Personal nu credem că acest al doilea val lexical dacic ar putea fi cît de cît însemnat, deoarece nu există motive ca dacii liberi, vorbitori ai unei limbi care devenise acum necunoscută populaţiei romanizate să exercite asupra latinei o influenţă mai profundă decît, de exemplu, goţii ale căror urme în vocabular le-au căutat pînă acum fără succes cîteva generaţii de lingvişti.

În ordinea descrescătoare a valorii raportului $\delta$, după cele două elemente constitutive iniţiale ale limbii române urmează, cu cifre apropiate (2,423 şi respectiv 2,244), cuvintele vechi slave şi cele cărora dicţionarul (e vorba de DEX) le dă o dublă explicaţie etimologică prin bulgară şi sîrbocroată. Diferenţiarea, cîtă este, dintre cele două



Scanned with OKEN Scanner

categorii a apărut numai după revizuirea etimologiilor în lumina DEX, căci potrivit informaţiilor din DLRM valorile $\delta$ erau practic identice: 2,41 pentru elementul vechi slav şi 2,42 pentru cuvintele cu dublă etimologie bg. & scr. Unităţile cedate de cea de a doua clasă în favoarea celei dintîi sînt în general posesoare ale unui număr relativ mare de derivate, ceea ce se reflectă în apariţia între productivităţile medii a unui decalaj, care ar mai putea încă să crească dacă s-ar include printre cuvintele vechi slave şi unii termeni slavi comuni dialectelor nord şi sud-dunărene, ca de exemplu *grădină* (8 derivate) sau *opări* (5 derivate). Oricum între valoarea $\delta \approx 3{,}9$ ce caracterizează fundamentul lexical primordial (fond latin + substrat) al limbii române şi raportul $\delta \approx 2{,}2 - 2{,}4$, propriu împrumuturilor slave comune, diferenţa rămîne apreciabilă.

Sensibil mai mică ($\delta = 1{,}933$) este productivitatea formaţiunilor onomatopeice. Creaţii spontane, deloc sau foarte slab condiţionate istoric, ele au putut lua naştere în orice vreme. Presupunînd că apariţia lor s-ar distribui uniform în timp ne-am putea aştepta ca vîrsta medie a acestora să se situeze în jurul a 50% din vechimea limbii române înseşi. Or, plasarea potrivit criteriului $\delta$ a acestor unităţi lexicale imediat în urma aportului vechi slav şi înaintea celorlalte influenţe *(bulgară, maghiară, turcă, neogreacă* etc.) corespunde tocmai unei atari datări.

Se cade să adăugăm totuşi că nici în cazul acestor elemente „motivate extralingvistic" nu se poate vorbi de o independenţă absolută în raport cu evoluţia contextului lingvistic în care au luat naştere. Există astfel cuvinte explicabile deopotrivă ca onomatopee şi ca împrumuturi din limbile vecine. Apropiat în DLRM de ucr. *mototolyty,* verbul *mototoli* este presupus a fi de origine onomatopeică de către autorii DEX. Pentru *mîţă* acelaşi dicţionar trimite la *alb. mica,* fără a exclude mai vechea ipoteză onomatopeică. Chiar şi în cazul unui diagnostic etimologic dat drept cert ca *scrîşni* ⟨ bg. *skărša, scr. skršiti* sau *şopot* ⟨ *v. sl. šiputu,* caracterul imitativ uşor perceptibil al cuvintelor româneşti lasă oricînd deschisă o portiţă pentru explicarea prin mijloace interne. De altfel, posibilitatea apariţiei independente în limbi diferite a unor onomatopee identice sau asemănătoare nu este deloc exclusă. Coincidenţa fonetică dintre *rom. miau* şi *fr. miaou* nu are nevoie de presupunerea vreunei influenţe reciproce, întrucît sunetele emise de organul vocal al pisicii constituie o explicaţie îndestulătoare a asemănării celor două cuvinte. Să mai subliniem şi că, la drept vorbind, nu există o incompatibilitate de principiu între geneza internă şi împrumut. Dimpotrivă, ele îşi pot da mîna (sub forma, deja unanim acceptată, a etimologiei multiple), concurînd la apariţia şi consolidarea unor familii de cuvinte de origine mixtă indi- şi alogenă. Aşa s-a întîm-



Scanned with OKEN Scanner

plat cu familia onomatopeei *trop (tropăi, tropăială, tropăit, tropăitură)*, îmbogăţită prin imigrarea *bg. scr. tropot*, împrumutat ca atare şi devenit apoi el însuşi generator de derivate *(tropoti, tropotire, tropotit)*.

Comparînd între ele tabelele 8.1. şi 8.3, se observă în cel de al doilea o creştere apreciabilă (cu 71 de unităţi) a numărului cuvintelor-bază de origine onomatopeică. Acest spor substanţial se datorează aproape în exclusivitate schimbării de optică intervenite în DEX faţă de DLRM în ceea ce priveşte înţelegerea conceptului de onomatopee. Pentru autorii primei lucrări nu sînt onomatopeice decît cuvintele cu ajutorul cărora se încearcă reproducerea unor sunete din natură *(balang, bîz, buf, fîl, mîr, oac, pleosc, sfor, tic-tac, vîj)*. Interjecţiile cu caracter neimitativ *(di, ehei, ha, hei, ho, phii, st)* figurează drept cuvinte de origine necunoscută. Nu aceeaşi este perspectiva DEX, care le asimilează şi pe cele din urmă cu onomatopeele, „clarificînd" dintr-un condei 71 de etimologii necunoscute. Rezolvarea ne pare cam superficială, avînd în vedere că interjecţiile se împrumută şi ele aidoma tuturor celorlalte cuvinte dintr-o limbă într-alta. Odată ce ştim că *haide* provine din turceşte, *raita* din maghiară, iar *vai*, poate, din latină, de ce am considera în bloc drept formate pe teren românesc toate interjecţiile a căror origine n-a fost încă identificată?

Pe scara productivităţii, imediat în urma formaţiunilor onomatopeice se află împrumuturile bulgăreşti. Pentru o situare corectă în diacronie a perioadei lor de pătrundere în limba română, este necesar ca în calculul potenţialului de derivare mediu al cuvintelor de origine bulgară să se ţină seama şi de cele 558 de unităţi (172 de cuvinte-bază şi 386 de derivate) cu etimologie dublă bulgaro-sîrbocroată, pe care simplul fapt de a fi imigrat pe două căi distincte nu le face cu nimic mai puţin bulgăreşti decît cele dintîi. Productivitatea medie ce se obţine procedînd astfel este $\delta = \frac{536 + 386}{287 + 172} = 2{,}009$, adică ceva mai mare decît cea a elementelor de provenienţă onomatopeică. Un calcul similar pentru aportul sîrbocroat ne conduce la $\delta = \frac{125 + 386}{167 + 172} = 1{,}507$, valoare simţitor mai mică, ce reflectă, în acord cu datele istorice, caracterul mai tîrziu al influenţei sîrbeşti în raport cu cea bulgărească.

În afara categoriilor deja discutate, singura componentă etimologică pentru care parametrul $\delta$ mai este supraunitar în tabelul 8.3 o constituie mulţimea cuvintelor de origine maghiară. Începută probabil în secolele XI—XII (absenţa elementelor maghiare din dialectele româneşti sud-dunărene fixează o clară limită *post quem* a acestei



Scanned with OKEN Scanner

influenţe), înrîurirea limbii maghiare asupra vocabularului românesc este oricum mai veche decît a celorlalte limbi neslave (turca, neogreaca, germana etc.). Cele mai însemnate aluviuni par să corespundă perioadei mai vechi a contactului româno-maghiar, căci majoritatea termenilor maghiari din vocabularul de bază figurau deja în primele texte româneşti cunoscute[1]. Limitîndu-ne la fondul principal lexical definit de Al. Graur, este de remarcat că peste trei sferturi (mai precis 77,4%) dintre cuvintele maghiare ce intră în alcătuirea sa posedă atestări documentare din secolele XV—XVI: *bănui, birui, cheltui, chibzui, chin, chip, făgădui, fel, gazdă, gînd, hotar, marfă (marhă), meşter, mîntui, neam, oraş, pildă, seamă (samă), sîrg, socoti, şoim*[2], *tăgădui, vamă, viclean (hitlean)*. Tot anterioară anului 1600 este şi consemnarea în scris a două dintre cele trei cuvinte din FPL a căror origine maghiară este incertă, dar destul de probabilă: *talpă (cf. magh. talp)* şi *tîlhar (cf. magh. talvaj)*. Vechimea aportului maghiar este aşadar apreciabilă şi dacă ceva pare ciudat în tabelul 8.3 aceasta nu este valoarea ridicată înregistrată de parametrul $\delta$ pentru împrumuturile maghiare, ci, tocmai dimpotrivă, faptul că $\delta$ este mult mai mic decît în cazul elementului bulgar (1,080 faţă de 1,868 sau chiar 2,009, dacă se ţine şi de cuvintele cu dublă etimologie bulgaro-sîrbocroată). Explicaţia probabilă a acestui decalaj rezidă, după opinia noastră, în ponderea însemnată acordată în DLRM regionalismelor şi arhaismelor de origine maghiară, termeni periferici, în majoritate neproductivi, a căror prezenţă diminuează într-o măsură exagerată numărul mediu de derivate pe unitatea lexicală primară (cuvînt-bază). Este suficient să menţionăm că, din totalul de 479 de cuvinte-bază împrumutate din ungureşte[3], 242, adică ceva mai mult decît jumătate,

---

[1] A. Rosetti (ILR p. 113—116 şi p. 328—330) dă lista acestora, dar din ea lipsesc cîteva cuvinte bogat atestate în textele din secolul al XVI-lea: *cheltui* (16), *mîntui* (133), *pildă* (170), *seamă (samă)* (22), *sîrg* (13), *socoti* (116). Cifrele din paranteze reprezintă frecvenţa absolută a cuvintelor respective în documentele anterioare anului 1600 stabilită de Claudia Tudose *(Vocabularul fundamental al limbii române vechi* în *Sistemele limbii* Ed. Academiei R.S.R., 1970, p. 119—164).

[2] Prima atestare la 15 ianuarie 1467 (G. Mihăilă: *„Dicţionar al limbii române vechi"* Ed. Enciclopedică, 1974, p. 163).

[3] La cele 453 de unităţi din tabelul 8.3 se mai adaugă 26 de cuvinte cu etimologie dublă magh. & scr. *(ceaptă, chinez 2, megies, pahar, rudă 2, sic 2, subă, toc)* magh. & tc *(surduc, tărîm)*, magh. & ucr. *(pelincă, rătez, tască)*, magh, & rus *(mondir, potroc)* magh. & pol. *(erhă, mustră)*, magh. & bg. *(podmol)*, magh, & sas *(ţandără)* şi multiplă: magh. & bg & scr *(satră)*, magh & scr & germ. *(păstîrnac)*, magh & rus & ucr. *(sîsîiac)*, magh. & pol & tc *(palancă)*.



Scanned with OKEN Scanner

sînt fie regionale *(almar, arădui, arneu, baie (mină), băcuiaţă, bănat, bărbînţă, bărnaci* etc.), fie învechite *(baştie, bir, căpeneag, ciocîrlat, conţ, honved, hotnog* etc.), fie şi una şi alta *(aldui, bărăţie, birău, borneu, jurui, nemeşug, palută, paşuş, ponoslui* etc.). Ar mai fi de adăugat la acestea alte cîteva cuvinte ce poartă menţiunea „maghiarism" *(făgădaş, potică, salău)* şi care, prezentînd, la rîndul lor, o arie de răspîndire limitată, pot fi asimilate, fără teamă de greşeală, regionalismelor[1]. Dacă facem abstracţie de toţi aceşti termeni, rămîn 234 de cuvinte-bază care au generat 426 de derivate, media productivităţii fiind $\delta = \frac{426}{234}$ = 1,821, valoare mult superioară celei anterior calculate. Acum capacitatea derivativă a materialului lexical maghiar s-a apropiat simţitor de a celui bulgăresc, dar cifrele nu mai sînt comparabile, căci împrumuturile bulgăreşti au fost considerate *tale-quale*, fără excluderea termenilor regionali şi învechiţi. Eliminîndu-i pe aceştia, raportul $\delta$ ajunge, pentru elementul bulgar, la 2,173 derivate per cuvînt-bază, dacă nu se iau în considerare cuvintele cu dublu etimon bulgaro-sîrbocroat, şi la 2,200 în cazul contrar. Aduse astfel la „acelaşi numitor", valorile corespunzătoare aporturilor maghiar şi bulgăresc s-au îndepărtat din nou una de alta, dar nici pe departe în măsura iniţială: de la 0,8—0,9 decalajul a scăzut la 0,3—0,4.

După cum se vede, valoarea parametrului $\delta$ depinde de ipoteza lexicologică adoptată. Rezultatele diferă în funcţie de luarea sau nu în considerare a unor termeni marginali (regionalisme şi/sau arhaisme), de includerea sau nu a cuvintelor cu etimologie dublă şi/sau multiplă etc. Posibilitatea determinării în mai multe moduri a productivităţii medii, cu rezultate uneori foarte diferite, complică bineînţeles problema găsirii unei corelaţii univoce între vechime şi capacitatea de generare de cuvinte noi. Valoarea parametrului $\delta$ depinde de prezenţa

---

[1] Observaţia, perfect valabilă pentru acest caz particular, nu poate fi generalizată, căci, de pildă, franţuzismele nu corespund unei zonări teritoriale, ci uneia sociale a limbii, ele sînt elemente de jargon. Ceea ce nu exclude existenţa regionalismelor de origine franceză: *baschină* (bluză de stambă) ⟨ fr. *basquine, întabula* (a înscrie un drept imobiliar în cartea funciară) ⟨ fr. *intabuler, parasol* (umbrelă de soare) ⟨ fr. *parasol, zoavă* (scurteică) ⟨ fr. *zouave;* după cum caracterul neologic al aportului galic nu este incompatibil nici cu perimarea, unele împrumuturi din franceză fiind deja resimţite drept desuete *(amoreză, convicţiune, demoazelă, soareá),* iar altele de-a dreptul arhaice: *fanal* (felinar mare) ⟨ fr. *fanal, pruă* (proră) ⟨ fr. *proue.*



Scanned with OKEN Scanner

în „cercul" lexical considerat a unui număr mai mare sau mai mic de elemente periferice (aici a fost vorba numai de termenii regionali sau învechiţi, dar similară este şi situaţia altor unităţi marginale ne- sau slab productive prin includerea ori eliminarea cărora productivitatea medie se vede sensibil afectată). La prima vedere, reducerea numărului total de cuvinte ar trebui să aibă drept consecinţă creşterea productivităţii medii. Lucrurile nu stau însă întotdeauna aşa, deoarece în procesul de sacrificare a termenilor periferici sînt antrenate şi derivatele, ceea ce face ca unele cuvinte-matcă să sufere însemnate „retrogradări" pe scara productivităţii[1]. Soarta indicatorului $\delta$ depinde atunci de jocul raporturilor numerice dintre cele două clase de unităţi susceptibile de a fi omise: clasa cuvintelor-bază şi cea a derivatelor.

Preferinţa unor autori de dicţionare explicative merge către includerea, cu precădere, a cît mai multor cuvinte-bază, pe temeiul observaţiei că acestea nu posedă, decît cu rare excepţii, o „formă internă", spre deosebire de derivate, al căror înţeles este uşor de dedus pentru cine cunoaşte sensul sufixului. Ne putem astfel dispensa de consemnarea în dicţionar a diminutivului *bordeiaş* odată ce îl avem pe *bordei,* dar sinonimul *argea* se cuvine neapărat explicat întrucît el nu aparţine unei familii cunoscute. Despuierea unor dicţionare axate preponderent pe termenii nederivaţi furnizează valori mai reduse ale productivităţii medii în raport cu alte lucrări lexicografice de volum similar, dar care acordă derivatelor un spaţiu mai însemnat. Este ceea ce se întîmplă, de exemplu, cu dicţionarul lui I. A. Candrea, de întindere asemănătoare cu DLRM, dar mai zgîrcit cu creaţiile lexicale româneşti. După D. Macrea el cuprinde 24 311 cuvinte moştenite sau împrumutate şi numai 18 958 de derivate şi variante, în timp ce în DLRM această a doua categorie, cu un total de 25 669 de unităţi, precumpăneşte asupra celei dintîi, cifrată la 23 504 cuvinte. Consecinţa acestei opţiuni este valoarea sistematic mai redusă a productivităţii

[1] Un singur exemplu: larg răspînditul *corn* ⟨ lat. *cornu,* cuvînt general românesc, numără printre derivatele sale atît regionalismele *cornări* = „a conduce plugul ţinîndu-l de coarne" şi *corneci,* diminutiv al termenului-matcă, cît şi arhaismele *cornărit* — „dare pe vitele cornute" şi *cornar* = „slujbaş care strîngea cornăritul". O selecţie mai severă poate înlătura din dicţionar aceste rude sărace ale lui *corn,* slăbindu-i poziţia în ierarhia unităţilor productive, fapt ce se repercutează, în cele din urmă, asupra productivităţii medii calculate a elementului latin.



Scanned with OKEN Scanner

medii pentru toate componentele etimologice ale vocabularului, ceea ce poate fi observat în tabelul 8.4 unde am alăturat, spre comparaţie, cele două şiruri de valori ale parametrului $\delta$, rezultate din statisticile întocmite de D. Macrea[1] pentru dicţionarele în discuţie.

*Tabelul 8.4*

| | LATIN | SUBSTRAT | SLAV | BULGAR | BG & SCR. | MAGH. | SCR. | TC. |
|---|---|---|---|---|---|---|---|---|
| Candrea | 3,24 | 3,61 | 2,09 | 1,21 | 1,27 | 1,25 | 0,65 | 0,69 |
| DLRM | 4,36 | 4,69 | 2,49 | 2,16 | 2,52 | 1,42 | 1,17 | 0,74 |

| | ONOM. | RUS. | LAT. SAV. | NESIGUR | NECUN. | GERM. | ITAL. | FRANCEZ |
|---|---|---|---|---|---|---|---|---|
| Candrea | 1,99 | 0,67 | 0,30 | 0,94 | 0,90 | 0,40 | 0,23 | 0,13 |
| DLRM | 3,07 | 0,79 | 0,61 | 1,57 | 0,92 | 0,57 | 0,98 | 0,50 |

Datele defalcate pe componente etimologice confirmă concluzia anterioară privind atenţia mai redusă acordată de I.A. Candrea termenilor derivaţi în raport cu cuvintele-bază. În plus, se observă că proporţiile în care sînt afectate diferitele aporturi variază apreciabil de la caz la caz. Productivitatea elementului turcesc rezultă sensibil aceeaşi din ambele dicţionare, pe cînd cea a împrumuturilor italiene şi franceze este de aproape patru ori mai mare în DLRM decît la I.A. Candrea. În aceste condiţii, este evident că încercarea de corelare a vechimii cu valoarea raportului $\delta$ e dinainte sortită eşecului.

Conştienţi de relativitatea indicaţiilor cronologice oferite de valorile medii ale productivităţii, nu putem totuşi să nu subliniem că, în mare, mai vechea observaţie a lui D. Macrea se confirmă. Indiferent de fluctuaţiile inerente datorate volumului şi calităţii diferite a corpusului lexical considerat, aporturile mai recente sînt caracterizate printr-o productivitate medie scăzută, valorile cele mai mici corespunzînd în ordine, potrivit tabelului 8.3, influenţelor: engleză (0,058), „franceză" (0,210), italiană (0,293), germană (0,379), rusă (0,398), latină savantă (0,462). Cea mai mare parte dintre acestea nu atinge nici jumătate din valoarea găsită pentru cel mai modest dintre aporturile timpurii discutate anterior. În intervalul valoric dintre

[1] După cum am mai arătat, avem motive să punem la îndoială exactitatea bilanţurilor întocmite de D. Macrea, dar, cel puţin în ceea ce priveşte dicţionarul lui I. A. Candrea, nu avem altceva mai bun de oferit. Dacă lingvistul a greşit în statisticile sale, să sperăm că a făcut-o în acelaşi sens în ambele cazuri.



Scanned with OKEN Scanner

contribuţiile lexicale vechi şi cele moderne se înscriu[1], în ordinea descrescătoare a productivităţii, aporturile neogrecesc (0,740), ucrainian (0,700) şi turcesc (0,489), acesta din urmă foarte apropiat de clasa împrumuturilor recente, semn că importul masiv de termeni turceşti trebuie situat în perioada contactului politic şi economic mai strîns cu Poarta din vremea domniilor fanariote. Deosebit de interesant şi demn de reţinut este faptul că, în medie, cuvintele de origine necunoscută nu sînt cu mult mai productive decît cele turceşti. Şansele ca ele să aparţină în majoritate unor straturi vechi ale limbii sînt deci reduse, ceea ce constituie un argument împotriva tendinţelor exagerate ale unor lingvişti (nu o dată improvizaţi) de a căuta în substrat sursa oricărui cuvînt de obîrşie încă neidentificată.

## 9. Clase de egală productivitate

După o primă luare de contact, prin intermediul imaginii globale oferite de ierarhia valorilor medii ale productivităţii, să facem un pas mai departe, trecînd la examinarea componenţei pe origini a listei celor mai productive cuvinte româneşti. Le-am enumerat în tabelul 9.1 pe primele 375, în ordinea descrescătoare a numărului de derivate pe care le au în DLRM. Cele de origine nelatină sînt notate diferit (cu caractere cursive) pentru a se pune în evidenţă creşterea densităţii lor pe măsura înaintării către zona productivităţilor mai modeste.

Avem în lista din tabelul 9.1 imaginea (inevitabil incompletă, dat fiind faptul că înşiruirea aici a tuturor celor mai bine de 25 000 de cuvinte nederivate din DLRM ar fi fost, din toate punctele de vedere, neconvenabilă), felului cum se organizează vocabularul pe criteriul productivităţii lexicale. Cuvintele-bază sînt grupate în clase de egală productivitate, cu atît mai bogate cu cît au generat mai puţine derivate. Primele 5 clase conţin cîte un singur element. Este vorba de cele mai prolifice cuvinte româneşti, care sînt substantivele *moş* şi *faţă*, adjectivele *bun* şi *alb* şi verbul *vedea*. Semnificativ e faptul că dintre

[1] Facem abstracţie de elementele săseşti şi ţigăneşti, insuficiente cantitativ pentru a permite o interpretare statistică satisfăcătoare.



Scanned with OKEN Scanner

*Tabelul 9.1*

| Clasă | Nr. de ordine | Cuvinte-matcă | Nr. de derivate |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 1 | 1 | *mos* | 38 |
| 2 | 2 | faţă | 35 |
| 3 | 3 | bun | 32 |
| 4 | 4 | vedea | 30 |
| 5 | 5 | alb | 25 |
| 6 | 6—11 | mînă, ochi, pădure, parte, *pic,* vînt | 23 |
| 7 | 12—15 | bate, pui, semn, verde | 22 |
| 8 | 16—20 | casă, *fărîmă,* roată, şarpe, trage | 21 |
| 9 | 21—26 | face, om, porumb, *prost,* pune, *vai* | 20 |
| 10 | 27—33 | coadă, *gol,* mărunt, scrie, ţară, *ţintă,* veni | 19 |
| 11 | 34—41 | frunză, galben, *grijă, moc, potrivă,* rece, simţi, spumă | 18 |
| 12 | 42—50 | cald, cunoaşte, frate, *lipi,* spăla, sta, *ţigan,* ţine, *vorbă* | 17 |
| 13 | 51—62 | *drag,* latură, mînca, *pod, pungă,* roşu, strînge, suci, sufla, umbla, usca, zi | 16 |
| 14 | 63—82 | arde, *băiat, clipi, colţ,* cruce, *dascăl,* drac, *groapă,* lucra, mamă, mare, mic, pămînt, piatră, *pînză,* tînăr, *tîrî, val,* zbura, zice | 15 |
| 15 | 83—100 | acru, aur, *copil,* cura, luci, negru, patru, *piti(c), plută,* putea, *rob,* scula, stea, şiră, şoarece, şterge, *voinic,* zece | 14 |
| 16 | 101—126 | bucată, corn, *crac,* da (vb.), doi, *drum,* fir, linge, lumina, minte, *murg,* neguţător, orb, pătură, prinde, *puscă,* rupe, *sfînt,* somn, *sticlă, scoală, soim,* tată, tăia, *trăsni,* urma | 13 |
| 17 | 127—162 | albastru, albi, alună, bătrîn, *brad,* cap, creţ, dulce, fată, floare, *gata, hoţ,* învăţa, masă, moale, *nisip,* nou, număra nun, *obraz,* pai, *pestriţ,* piele, *prund,* sămînţă, schimba, scoarţă, *sold, tăvăli, tipar, topi,* trece, *turc,* urî, viu, zăcea | 12 |
| 18 | 163—212 | *arendă, balang, boier, butuc, chiabur,* cerc, cere, cufunda, domn, *fîs, fricǎ,* fum, furcă, *gînd,* întinde, lung, mătură, mult, naşte, *nene,* noapte, nume, oaste, pană, pas, păgîn, pescar, *petic,* piept, pisa, *pîndi, pleată, plug, prieten,* puţin, salcie, sat, sîngera, *socoti,* strîmt, suflet, suge, *surub, talpă,* tare, *ţelină,* uşor, vierme, vînăt, *zur* | 11 |



Scanned with OKEN Scanner

*(continuare) Tabelul 9.1*

| 0 | 1 | 2 | 3 |
|---|---|---|---|
| 19 | 213—286 | argint, armă, *babă, bale, baltă,* barbă, *băl, buf,* cădea, călca, *cioc, clăti,* crede, dator, deschide, drept, *dura,* fecior, *flăcău,* gros, împărţi, încărca, întoarce, întreg, înţelege, *jupui, lanţ,* lat, legăna, lemn, limpede, lin, *mîndru, mladă, mulţumi,* munte, nor, *noroc,* păstor, păun, *pis,* plin, porc, putred, *raz,* rîncheza, român, *rumen, săpun,* săpa, sărac, scaun, *sfat, sigur,* spin, supune, şa, ştie, *tabac,* tămîie, *tătar, tîlhar, tîrn,* toca, tremura, *ţeapă,* urs, vecin, *vesel,* vin, *vîj, vîrf,* vulpe, vultur | 10 |
| 20 | 287—375 | arc, *bold, brazdă, buche,* car, carne, călare, cenuşă, cetate, *cioban,* ciocan, ciur, *cîrciumă, clamp,* coardă, *colac,* cot, creşte, *crai, croi, cucoană,* cugeta, *cumpănă, curma, dusman,* ferice, fierbe, fiinţă, *fluier, fluture,* foaie, foc, fund, găină, *ham, hotar, împrejur, învîrti,* juca, junghia, lingură, marmură, *mînz,* muced, muia, *năravi,* neted, *opri, pahar,* pasăre, *păcală, pită,* plînge, *pofti,* popor, *potcoavă, poveste, puf* (interj.), ram, rotund, sac, *săvîrsi, scîrţ,* seca, *slugar (nic),* soră, sterp, *străin, taină, tearfă,* toamnă, *troian, trup,* tufă, turna, *ţarc,* ţese, *ţol,* uita, umbră, *ursi,* urzică, vale, viaţă, *vifor,* voi, *zvîc* | 9 |

aceste 5 cuvinte, 4 sînt de origine latină şi unul, cel mai productiv, aparţine substratului. În clasa următoare apar de asemenea 5 termeni latini moşteniţi, la care se adaugă un cuvînt onomatopeic: *pic*. Izolarea acestuia în raport cu celelalte formaţiuni de acelaşi fel (onomatopeele imediat următoare în ordinea rangului apar abia în clasa a 18-a şi nu au nici pe jumătate atîtea derivate ca *pic*) ar putea fi pusă în legătură cu marea sa vechime. Nu ni se pare deloc exclusă prezenţa sa încă din latina dunăreană, cu atît mai mult cu cît nimic din forma cuvîntului nu se opune unei atare datări. E un aspect care merită reţinut întrucît nu ştim ca pînă acum lingviştii să-şi fi pus, de pildă, problema dacă nu există cumva în româneşte şi onomatopee provenite din substrat aşa cum există, cu certitudine, unele împrumutate mai recent din limbile vecine.

Începînd cu cea de a 7-a, alcătuită numai din elemente latine, clasele sunt din ce în ce mai bogate. Numărul de cuvinte pe care le



Scanned with OKEN Scanner

conţin este în genere[1] crescător, ceea ce se poate urmări în tabelul 9.2., unde este înscris şi numărul de membri ai celor 9 clase absente,

*Tabelul 9.2.*

| Clasa | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. de cuvinte | 4 | 5 | 6 | 7 | 8 | 9 | 12 | 20 | 18 | 26 | 36 | 50 |

| Clasa | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. de cuvinte | 74 | 89 | 122 | 162 | 270 | 321 | 478 | 690 | 1357 | 2716 | 19181 |

din pricina bogăţiei lor, din tabelul 9.1. În primele 20 de clase din cele 29 pe care le consemnează tabelul 9.1., mai mult de jumătate dintre cuvinte sînt de origine latină: adăugînd la cei 228 de termeni moşteniţi două unităţi formate în româneşte din material lexical de aceeaşi proveniență *(mulţumi* şi *împrejur),* rezultă că fondul latin reprezintă 61,9% din total. Dacă se iau însă în considerare numai primele 10 clase, acest procent creşte la 78,8%. În zona celor mai productive cuvinte româneşti, elementul latin este aşadar precumpănitor. În schimb, din totalul cuvintelor productive din DLRM, numai 20,2% sînt moştenite din latineşte. Ierarhizarea unităţilor lexicale după numărul de derivate pe care le-au generat confirmă astfel rolul hotărîtor al fondului latin în constituirea vocabularului românei actuale, conducînd la rezultate care seamănă calitativ cu cele furnizate de inventarele de frecvenţă. Şi în cazul acelora prezenţa latină era net dominantă în zona rangurilor mici, chiar dacă în cifre absolute împrumuturile ulterioare depăşeau per total componenta latină.

Examinarea tabelului 9.1. ne prilejuieşte şi constatarea că, deşi important, indicatorul productivitate nu este îndestulător pentru stabilirea însemnătăţii cuvintelor. Există termeni aflaţi în declin sau chiar ieşiţi cu totul din uz ce au generat familii bogate, ai căror membri sînt adesea mai cunoscuţi decît cuvîntul-matcă. Iată cîţiva dintre ei:

*moc* (et. nec.) cu derivatele: *moacă, mocan, mocancă, mocăi, mocăială, mocăit, mocănaş, mocăncuţă, mocăneasă, mocănesc, mocăneşte, mocănime, mocîrţă, mocîrţan, mocofan, mocăniţă, măgădan, măgădău*

[1] Excepţie face numai clasa a 14-a, cu puţin mai cuprinzătoare decît cea imediat următoare.



Scanned with OKEN Scanner

*pit* (cf. slav) cu derivatele: *pitic, pitcoace, piti, piticenie, pitici, piticire, pitiş, pitit, pitula, pitulat, pitulice, pituliş, tupila, tupiluş*

*băl* (⟨ v. sl. *belu*) cu derivatele: *bălai, bălan, bălăiel, bălăior, bălănel, bălănuţ, bălăoară, bălăuc, băluşcă, băluţ*

*raz* (⟨ scr. *raz*) cu derivatele: *răzui, răzălui, răzăluială, răzăluire, răzăluit, răzuială, răzuire, răzuitoare, răzuitor, răzuş*

*tabac* (cu sensul de *tăbăcar;* ⟨ tc. *tabak*) cu derivatele: *tăbăcar, tăbăcăreasă, tăbăcăresc, tăbăcărie, tăbăceală, tăbăci, tăbăcire, tăbăcit* (subst.), *tăbăcit* (adj.), *tăbăcitor*[1]

*slugar* (⟨ bg. *slugar*) cu derivatele: *slugarnic, slugăreală, slugăresc, slugăreşte, slugări, slugărie, slugărime, slugărit, slugărnici, slugărnicie*

Evident, nici unul dintre aceste cuvinte nu (mai) poate fi socotit important la ora actuală. A le reţine într-un proiect de vocabular fundamental pe considerentul productivităţii ar fi o greşeală, pe care Al. Graur s-a ferit să o facă atunci cînd a propus variantele sale de fond pricipal lexical. Ţinînd mereu seama de uzaj, deşi nu s-a bazat pe statistici, acesta a evitat în general să absolutizeze criteriul productivităţii, iar atunci cînd un derivat a ajuns să fie mai folosit decît ascendentul său nu a ezitat să-l reţină în locul acestuia ca reprezentant al familiei respective de cuvinte. Aşa se face că în FPL, sînt prezente *ramură,* în locul mai vechiului *ram* din care provine, şi *leagăn,* derivat regresiv de la absentul *legăna.*

În treacăt fie spus, noţiunea de familie de cuvinte pare să fi fost folosită de Al. Graur într-un sens mult mai larg decît am făcut-o noi. Cuvinte cu formă asemănătoare şi înţeles înrudit, deşi nederivate *în româneşte* unul din altul, au fost considerate drept membri ai aceleiaşi familii. Asimilarea s-a bazat fireşte pe existenţa unui etimon comun mai îndepărtat într-una dintre limbile de origine[2]. Al. Graur nu o spune explicit, dar dacă a inclus în FPL neologismul *garant* a făcut-o fără îndoială nu pentru importanţa în sine a acestui cuvînt, ci pentru că el reprezintă o rădăcină care a generat (dar nu în sînul limbii române) pe *garanta* şi *garanţie.* Neproductivul *negoţ* (⟨lat. *negotium*), monosemantic şi el ca şi *garant,* îşi datorează probabil prezenţa în fondul principal calităţii de exponent al unei „familii" în sens larg, ce cuprinde şi pe foarte prolificul *neguţător* (cu derivate ca *negustor,*

---

[1] Din varianta *tărbăci* provine şi derivatul regresiv *tărbacă.*

[2] Altminteri nici un lingvist nu le-ar fi grupat în aceeaşi familie, cum nu pot fi grupate la un loc *sprint* şi *sprinten* care seamănă, totuşi, atît formal cît şi semantic!



Scanned with OKEN Scanner

*negustoraş, negustorească, negustoresc, negustoreşte, negustori, negustorie, negustorime, negustoriţă, negustoraş, neguţătorească, neguţători, neguţătorie*), care provine însă dintr-un etimon diferit (lat. *negotiatorius*). Din punctul său de vedere, procedeul era corect, căci Al. Graur nu îşi propunea să studieze productivitatea cuvintelor, ci numai să asigure reprezentarea în al său FPL a principalelor „cuiburi" lexicale româneşti. Inconsecvent este numai atunci cînd, după ce declară că a acordat „fiecărei familii un singur loc" reţine în listele sale pe *întuneric* alături de *întuneca*[1]. În schimb, lasă pe dinafară foarte importantul cuvînt *suflet*, probabil în virtutea includerii lui *sufla*, descendent din sursa latină mai îndepărtată a celui dintîi: *suflet* ⟨lat. *sufflitus* ⟨lat. *sufflare* ⟩ rom. *sufla*. Pentru vorbitorul român din zilele noastre legătura dintre *întuneca* şi *întuneric* este însă mult mai evidentă decît cea dintre *suflet* şi *sufla*.

Dincolo de asemenea obiecţii de detaliu, nu este mai puţin adevărat că, în general, principiul lui Al. Graur funcţionează, în sensul că, în conştiinţa vorbitorului nespecialist, cuvinte de provenienţe atît de diferite ca, de exemplu, *locui* (⟨magh. *lakni*), *localnic* (⟨ rus. *lokal'nyj*), *local* (⟨ fr. *local*), *loco* (⟨lat. sav. *loco*) aparţin toate familiei lui *loc* (⟨lat. *locus*). Fără a deriva direct din el, acestea contribuie la consolidarea poziţiei lui *loc* în vocabularul fundamental, la fel cum cuvintele *bărbier* (⟨ ngr. *barberis*), *barbetă* (⟨ it. *barbetta*) şi *barbişon* (⟨ fr. *barbichon*) îl sprijină pe urmaşul direct al etimonului lor de a doua mînă care este *barbă* (⟨ lat. *barba*).

Pe noi, o procedare analoagă ne-ar fi condus la o ordine de importanţă a cuvintelor mult diferită de cea din tabelul 9.1. Adjectivul *alb* (⟨ lat. *albus;* 25 de derivate), susţinut de substantivul omonim (⟨ lat. *albus*), de verbul *albi* (⟨ lat. *albire;* 12 derivate), dar şi de neologisme ca *albinism* (⟨ fr. *albinisme*), *albinos* (⟨ fr. *albinos*) etc. ar ajunge în capul listei ca *pater familias* al celui mai întins clan lexical românesc, după cum mai modestul *sare* (⟨ lat. *sal;* 5 derivate) ar promova spectaculos în ierarhie cu concursul unor *săra* (⟨ lat. *salare*); 10 derivate), *saramură* (⟨ ngr. *salamura*; 2 derivate) şi poate chiar al neologismelor *salin, salină, salinitate* etc. Scrupulul de a nu amesteca criteriile şi dorinţa exploatării pînă la capăt a sugestiilor pe care le oferă fenomenul derivării pentru o caracterizare sistemică a vocabularului ne-au oprit de la asemenea grupări, care, în măsura în

[1] Şi aici etimoanele sînt probabil diferite (*întuneca* ⟨lat. *in-tunicare* şi *întuneric* ⟨rom. *în + tuneric* ⟨lat. *tenebricus*), dar ascendenţa divergentă nu este perceptibilă pentru neiniţiat.



Scanned with OKEN Scanner

care fac să intervină unii factori subiectivi (pentru ce categorii de vorbitori este sesizabilă înrudirea cutăror cuvinte?), se află în dezacord cu tipul de demers pentru care am optat.

## 10. Dependenţa de origine a puterii generative a cuvintelor

Să examinăm mai îndeaproape comportarea sub raportul productivităţii a cuvintelor de diferite origini. Începem, cum e şi firesc, cu fundamentul lexical al limbii române care este:

### 10.1. Fondul latin moştenit

Dintre cele 1802 cuvinte latine conservate de limba română pe care le cuprinde Dicţionarul Limbii Române Moderne, 1328, adică aproape trei pătrimi (73,70%), au generat derivate pe teren românesc. Restul de 474 (26,30%) sînt neproductive[1]. Ne referim exclusiv la derivarea propriu-zisă, deoarece compusele, atît cele alcătuite numai din material latin moştenit cît şi cele de origini mixte au parte de o tratare separată. Avînd în vedere însemnătatea primordială a elementului latin, dăm mai jos, în ordinea descrescătoare a numărului de derivate, lista celor 1328 de unităţi productive:

| | |
|---|---|
| 1. faţă | 35 de derivate |
| 2. bun | 32 de derivate |
| 3. vedea | 30 de derivate |

[1] La nivelul selecţiei operate de DLRM. Aici şi în continuare, cînd vorbim de sterilitate lexicală o înţelegem în sens relativ, raportată la conţinutul dicţionarelor utilizate. Nu putem afirma că un cuvînt este absolut neproductiv cît timp nu posedăm un dicţionar tezaur încheiat al limbii române care să ne dea certitudinea că unitatea respectivă nu a dat naştere nici unui derivat. În plus, trebuie ţinut seama de posibilitatea ca unele derivate să se fi pierdut înainte de a fi fost consemnate în scris.



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 4. | alb | 25 de derivate |
| 5-9. | mînă, ochi, parte, pădure, vînt | 23 de derivate |
| 10-13. | bate, pui, semne, verde | 22 de derivate |
| 14-17. | casă, roată, şarpe, trage | 21 de derivate |
| 18-21. | face, om, porumb, pune | 20 de derivate |
| 22-26. | coadă, mărunt, scrie, ţară, veni | 19 derivate |
| 27-31. | frunză, galben, rece, simţi, spumă | 18 derivate |
| 32-37. | cald, cunoaşte, frate, spăla, sta, ţine | 17 derivate |
| 38-45. | latură, mînca, roşu, strînge, sufla, umbla, usca, zi | 16 derivate |
| 46-57. | arde, cruce, drac, lucra, mamă, mare, mic, pămînt, piatră, tînăr, zbura, zice | 15 derivate |
| 58-70. | acru, aur, cura, luci, negru, patru, putea, scula, stea, şină, şoarece, şterge, zece | 14 derivate |
| 71-87. | bucată, corn, da (vb.), doi, fir, linge, lumină, minte, neguţător, orb, pătură, prinde, rupe, somn, tată, tăia, urma | 13 derivate |
| 88-111. | albastru, albi, alună, bătrîn, cap, creţ, dulce, fată, floare, învăţa, masă, moale, nou, număra, nun, pai, piele, sămînţă, schimba, scoarţă, trece, urî, viu, zăcea | 12 derivate |
| 112-142. | cerc, cere, cufunda, domn, fum, furcă, întinde, lung, mătură, mult, naşte, noapte, nume, oaste, pană, pas, păgîn, pescar, piept, pisa, puţin, salcie, sat, sîngera, strîmt, suflet, suge, tare, uşor, vierme, vînăt | 11 derivate |
| 143-187. | argint, armă, bale, barbă, cădea, călca, crede, dator, deschide, drept, fecior, gros, împărţi, încărca, întoarce, întreg, înţelege, lat, legăna, lemn, limpede, lin, munte, nor, păstor, păun, plin, porc, putred, rîncheza, român (subst.), săra, scaun, spin, supune, şa, şti, tămîie, toca, tremur, urs, vecin, vin, vulpe, vultur | 10 derivate |
| 188-230. | arc, car, carne, călare, cenuşă, cetate, ciur, coardă, cot, creşte, cugeta, ferice, fierbe, fiinţă, foaie, foc, fund, găină, juca, junghia, lingură, marmură, muced, muia, neted, pasăre, plînge, ram, rotund, sac, seca, soră, sterp, toamnă, tufă, turna, ţese, uita, umbră, urzică, vale, viaţă, voi (vb.) | 9 derivate |



Scanned with OKEN Scanner

231-286. adînc, albină, amesteca, apoi, aramă, ascuţi, broască, brumă, buciuma, căra, cîine, coace, credinţă, cuib, despărţi, dor, Dumnezeu, încet, închega, jos, jug, june, lînă, margine, măcelar, măsura, mulge, must, mut, oală, păpuşă, pereche, peşte, pieri, preţ, primăvară, rar, rădăcină, răsări, rău, sălbatic, săpa, scut, soţ, spic, spurca, stoarce, strica, toarce, urzi, ustura, vargă, vas, văduv, vărsa, vindeca — 8 derivate

287-374. acoperi, adormi, ajunge, alege, amăgi, amorţi, aspru, avea, bour, bulbuc, cal, cale, ceară, cîmp, coase, cuceri, cuprinde, curte, dinte, dos, drept, faur, fier, frige, frînge, fulgera, greu, iarbă, iepure, intra, înălţa, încheia, închide, încurca, înota, însul, judeca, lega, leşie, lup, merge, mijloc, mişel, mişca, muiere, nas, negură, nuntă, oaie, orz, ospăta, pace, păcurar, păr, părinte, păţi, peţi, pieptăna, poamă, rade, rîu, rugină, sarcină, sătul, scălda, scărmăna, scînteie, scurt, scutura, seară, secera, semăna, serba, sete, singur, sperla, suferi, şuiera, tot, turtă, umed, uşă, văr, vechi, veghea, vergea, zeamă — 7 derivate

375-505. acăţa, adevăr, afuma, amar, apăra, apleca, aprinde, apuca, aripă, astîmpăr, astupa, asupra, aşeza, auzi, berbec, blînd, boci, braţ, brîncă, bute, caier, căpăţînă, cărunt, căuta, cătuşă, chema, cîştiga, coastă, creastă, creştin, cuc, culege, cumpăra, curge, cuţit, despica, deştepta, drege, durea, fapt, fereca, fi, fierbinte, flacără, floc, frămînta, frumos, fura, gaură, ger, gusta, guşă, inel, inimă, împărat, împlini, înainte, îndărăt, îndesa, înflori, înfunda, îngrăşa, înnoda, însura, învia, junc, lac, lacrimă, lamură, lăsa, lege, legumă, loc, luntre, maţ, măcina, măciucă, mărăcine, mierlă, minune, mişuna, mîna, morar, muri, muscă, muşca, nimic, nuc, păcat,



Scanned with OKEN Scanner

pătrunde, piedică, pişca, ploaie, plumb, pom, prun, pumn, purta, rămîne, repede, revărsa, roade, rost, ruşine, sălta, sănătos, sări, scădea, scîndură, soartă, sorbi, sparge, spulbera, stinge, striga, stup, sui, surpa, surzi, trepăda, trimite, turna, ţărm, ucide, umfla, vară, văcar, vătăma, vie, vînător, zestre — 6 derivate

506-621. aciua, acolo, aduce, aduna, aiurea, ajutor, alinta, aluneca, ameţi, aşchie, atinge, bărbat 1, bărbat 2, bea, capră, cărare, cerca, chingă, cinci, cînepă, corn, crăpa, crunt, cui, cuviinţă, cuvînta, depăna, deşert, doagă, duios, dumica, fermeca, fior, flămînd, fragă, frig, frînghie, fugă, fur, greaţă, grec, gură, ierta, ieşi, in, îmbina, împunge, înalt, încălţa, îndura, îneca, înfrîna, înger, înşela, întuneca, joc, jura, lapte, lăuda, lăcrima, leşina, lua, lupta, măiestru, măr, mărita, miel, mîţă, mîngîia, mormînt, muc, palmă, păduche, păsa, pierde, pîntece, plăcintă, poartă, pregeta, preot, priveghea, prinzi, răpi, rătăci, ridica, roşu, sare, scăpăta, scuipa, scurge, soare, spinteca, strîmb, subţia, şedea, teme, tîrziu, treiera, trist, tu, uda, umple, undă, unge, unt, urca, ureche, urla, varză, veşmînt, viespe, vinde, vis, vîntura, zeu, zgîria — 5 derivate

622-769. ac, acela 1, acest, aiest, alerga, alina, apă, apune, arunca, ascunde, asemăna, aştepta, baie, baieră, băşică, biserică, broatec, bucă, carpen, carte, cămaşă, căsca, căţel, certa, ceaţă, cerceta, cireaşă, cînta, cîrnat, codru, corb, crud, cunună, curea, cuscru, cuveni, dărăpăna, dărîma, des, descăleca, descărca, descoperi, detuna, dezlega, dormi, duce, dumesnic, falcă, făta, femeie, fiu, fîntînă, freamăt, freca, găună, gheb, ghem, gras, îmbăta, împăca, înăuntru, încinge, încuia, înfăşura, îngheţa, înjuga,



Scanned with OKEN Scanner

înţelept, lătra, lepădа, lînced, lumînare, mie, minciună, mînie, muta, ninge, oaspete, ou, paltin, papură, par, păcură, părea, pepene, petrece, picior, piuă, plai, plăcea, pleca, plop, porumbă, pradă, prelung, pricepe, pupa, purice, putină, rări, răşchia, răzbuna, rîde, rouă, rumega, saţ, săgeată, săgeta, sătura, scară, scorbură, scurma, scurta, seceră, sfîşia, solz, spată, spăimînta, spînzura, sprînceană, stîng, strănuta, strecura, strigă, stuf, subţire, suna, supăra, sus, suspina, tăciune, tei, trei, trufă, tulbura, tulbure, tunde, turba, tuşina, umăr, unghie, uni, ura, vădan, venin, vioară, vîna, vîrtos, volbură — 4 derivate

770-924. abate, acela 2, adăpa, adăuga, afîna, afla, ajuta, albie, alboare, amărî, apăsa, apropia, arătа, arbore, aromân, asculta, asuda, aşterne, aţă, aţipi, aţîţa, bătaie, blestema, bou, căldare, căprior, cercel, cerne; ciumă, culca, cumpăta, cuteza, cutremura, defăima, degera, deget, dejuga, descălţa, desface, dezbina, dezbrăca, dezghioca, dezmierda, dezvolta, doamnă, faşă, feri, fîn, foamete, fraged, frunte, fugi, fumega, geamăn, ginere, greaţă, greier, iarnă, ied, iţă, îmbrăca, îmbuca, îmbuiba, împila, împrumut, încerca, închina, îndupleca, înfige, înfoia, înfrînge, îngăla, înghiţi, îngîna, îngreuna, îngroşa, îngurzi, întîi, întortochea, învinge, înviora, la (prep.), lărgi, limbă, lume, luneca, lut, mănunchi, mărgea, măsea, mătuşă, mesteca 1, miez, mîine, moină, mort, mugi, mur (bot.), mura, mustaţă, neg, nepot, nimic, nod, olar, orbeca, os, paşte (vb.), pată, păşune, pătul, pedestru, pieptene, pîine, pulpă, pupăză, putinţă, putoare, rază, răposa, răsuna, răşină, rîma, rînced, roura, ruga, săruta, scărpina, scoate,



Scanned with OKEN Scanner

sfîrtica, smulge, spune, spuză, stîmpăra, sturz, sufleca, sughiţa, sumete, sunet, surcel, şale, şerb, taur, tăcea, tîmplă, trunchi, tun, turmă, ţîţă, ulcea, unchi, veşted, vintre, vîlcea, zbiera — 3 derivate

925-1105. acum, adăpost, afunda, ager, ai, alunga, ameninţa, amiază, amuţi, anina, ara, armăsar, asurzi, atîta, (adv.), atrage, bine (adv.), boteza, bucea, bunătate, caş, căldură, căpăstru, căpăta, cărbune, cearcăn, cerb, ceteră, cheie, cîntec, cît (adv.), coamă, comînda, conteni, copt, creier, culme, cum, cumineca, cumnat, cupă, cuptor, cutare, cutreiera, degetar, deprinde, descînta, desfăşura, despuia, deşira, deţine, dezgheţa, dezmetici, dihoca, dura, fagure, faptă, făină, fereastră, fierar, foarfece, foi, fraged, frasin, frîu, frupt, fuma, furnica, furnică, fus, fuscel, geme, genunchi, grăunţă, grîne, grîu, holba, iapă, împinge, împresura, împuţi, încăleca, încăpea, înceta, încinge (term.), înconjura, încotro, îndemna, înfăşa, înfrunta, îngemăna, îngenunchea, îngîmfa, îngusta, înjura, însenina, înspăimînta, înşela, întărîta, înteţi, întîmpla, întreba, întunerici, înţesa, învenina, jneapăn, judeţ, larg (adj.), lăptar, limbut, lung (subst.), mai (subst.), marţ, mei, merinde, mesteacăn, minţi, mioară, mior, mîneca, mînecă, mustra, muşuroi, negru, niscai, nouă, nuia, opt, oua, păr (bot.), pleca („porni"), poimîine, prăda, pulbere, purcede, purcel, puroi, puţ, rage, răpune, răspunde, răşchira, repede, rîndunea, român (adj.), sapă, scoace, searbăd, secară, semăna (⟨ lat. seminare), senin (adj.), spată (osteo.), spinare, spîn, strămuta, străpunge, strepezi, stupi, şapte, şase, şchiopăta, topsica, trifoi, turturea, tuşi, ţînţar, uita („privi"), urcior 1, urcior 2, urdina, vacă, vergură, vită, vitreg, viţă, viţea, viţel, vîltoare, vînă, zgaibă, zmăcina, zvînta — 2 derivate



Scanned with OKEN Scanner

1106-1328. abia, adăsta, adia, adică, agest, ajuna, altul, amnar, amu, an (adv.), apărea, aproape, arete, arici, arină, arţar, asmuţi, astruca, atîta (pron.), bată, bălţat, băutură, beat, beţiv, brăţară, bucium, buged, burete, car (zool.), călcîi, cătină, căţea, căuş, ce, ceapă, cel, celar, cep, cer, cerbice, cheotoare, ciutură, cît (prep.), cîte, coif, coraslă, coş (⟨ lat. cossus), crăpătură, cucui, cumnată, cununa, cuvînt, deda, deretica, descheia, descinge, descuia, desculţa, desfăca, desfăşa, desfira, deturna, dimineaţă, duminică, el, eu, falce, făurar (⟨ lat. febr(u)arius), făget, fecioară, ficat, fie (⟨ lat. filia), fiu, fin, foale, forfeca, friptură, frînc, frumos, funie, fulger, geană, ghindă, ghioagă, greutate, gruie, ie, ienupăr, ierna, im, ist, împletici, împlînta, încă, încălţa, încoace, încumeta, încura, înfiera, înfrunzi, înfuleca, îngust, înspica, întîi, întuneric, înţina, înveşte, învolt, judecată, juncă, jur, larg (subst.), laţ, lăcustă, legătură, limbric, lindină, linte, lună, martor, marţi, mădular, mătreaţă, meserie, mia, miere, mieriu, mira, mistreţ, mînea, moară, muc, muchie, muşchi 1, muşchi 2, naie, nărtos, nea, negru (subst.), nepoată, nişte, niţel, noră, număr, oară 1 (⟨ lat. ovaria), oară 2 (⟨ lat. hora), papă (⟨ lat. pappa), pară, (bot.), pască, paşti, păros, pecingine, perete, piersică, ploua, pomet, prînz, pucios, punte, purcea, puşchea, ridiche, rîie, rîpă, rîs (⟨ lat. risus), roib, rug (bot.), rugi, salbă, săcui, săptămînă, scăpa, scroafă, sec, secetă, secure, sine, sîmcea, sîn, sînge, sînt, soacră, socru, sorb (bot.), spaimă, spate, spre, strat, strămutare, strup, sub, sulă, sumuţa, surd, şchiop, şes (subs.), timp, timpuriu, tindă, tindeche, toartă, tort, treaptă, tuna, tuse, ţeastă,



Scanned with OKEN Scanner

ţest, ud (adj.), un, unda, ungher, unghi, untură, unu, urdoare, vad, viermănos, viers, vineri, vioară (muz.), vînăt (adj.), zbate, zînă — 1 derivat

Trecem, nu fără părere de rău, peste multiplele sugestii (analiza distribuţiei cuvintelor pe categorii morfologice, clase semantice etc.) pe care ni le oferă acest material lexical (enumerarea completă a componentei productive a fondului latin din DLRM constituie în acest sens o invitaţie implicită şi o bază de pornire pentru alte cercetări) şi venim la problema ce ne preocupă: legea de repartiţie a numărului de cuvinte de origine dată (latină, în cazul de faţă) în funcţie de nivelul lor de productivitate. Din lista de mai sus, rezultă că 1328 de unităţi lexicale au cel puţin cîte 1 derivat, 1105 — cel puţin 2, 924 — cel puţin 3, 769 — cel puţin 4, .621 — cel puţin 5 etc. Tabelul 10.1. sintetizează aceste cifre, cu menţionarea şi a logaritmului natural al numărului de cuvinte respectiv, avînd în vedere că reprezentările grafice se vor face şi în continuare în scară logaritmică.

*Tabelul 10.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv-bază | 1328 | 1105 | 924 | 769 | 621 | 505 | 374 | 286 | 230 | 187 | 142 |
| ln (Nr. cuv.-bază) | 7,19 | 7,01 | 6,83 | 6,64 | 6,43 | 6,22 | 5,92 | 5,65 | 5,44 | 5,23 | 4,95 |

| NR. DERIVATE ≥ | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 111 | 87 | 70 | 57 | 45 | 37 | 31 | 26 | 21 | 17 | 13 |
| ln (Nr. cuv.-bază) | 4,71 | 4,47 | 4,25 | 4,04 | 3,81 | 3,61 | 3,43 | 3,26 | 3,04 | 2,83 | 2,56 |

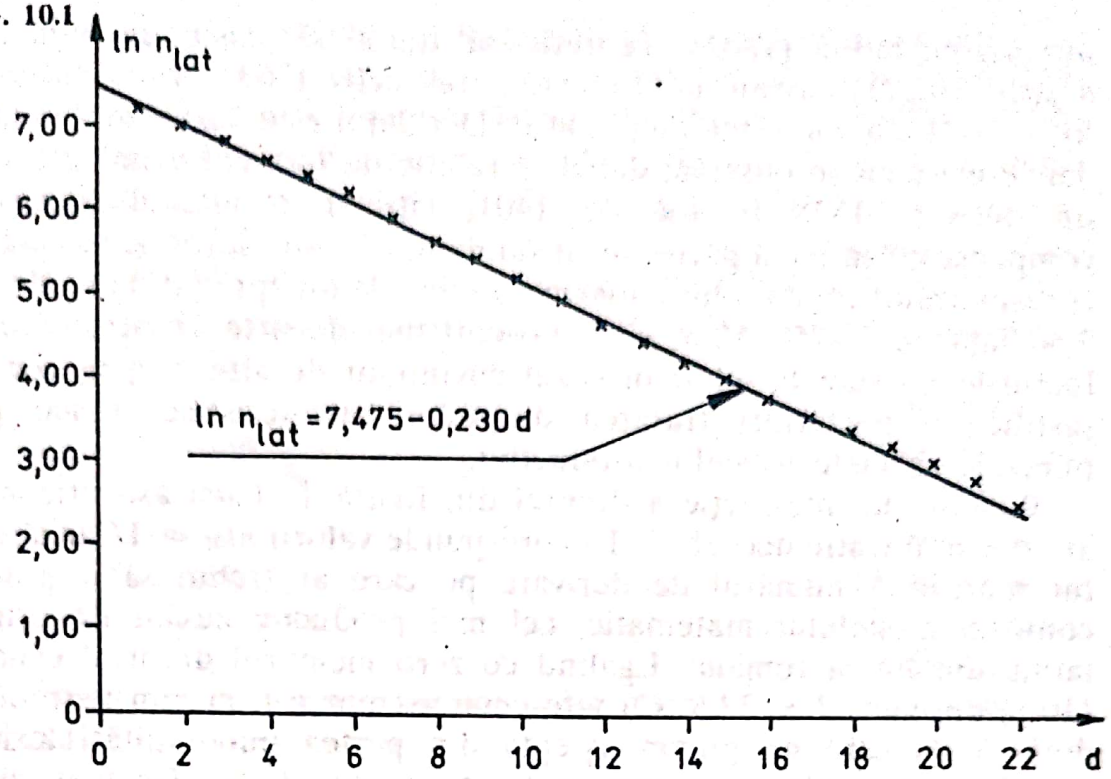
După cum se vede, ne-am limitat la clasele cu cel puţin 10 membri, întrucît coborîrea la valori mai mici o socotim insuficient de relevantă sub raport statistic.

Reprezentarea grafică a cifrelor din tabelul 10.1. într-un sistem cartezian avînd în abscisă numărul de derivate (linia 1 a tabelului) şi în ordonată logaritmul numărului de cuvinte-bază (linia 3) ne pune în faţa unei mulţimi de 22 de puncte dispuse aproape perfect coliniar (fig. 10.1.). Rezultatul este deosebit, iar implicaţiile lui se vor dovedi în continuare importante pentru descrierea în termeni cantitativi a



Scanned with OKEN Scanner

FIG. 10.1




fizionomiei vocabularului. Cu ajutorul metodei „celor mai mici pătrate"[1] se pot determina parametrii dreptei ce trece prin (printre) aceste puncte. Ecuaţia rezultantă este:

$$ln\, n_{lat} = 7{,}475 - 0{,}230\, d \qquad (10.1)$$

unde $d$ e numărul de derivate al unui cuvînt, iar $n_{lat}$ — numărul de cuvinte de origine latină care au cel puţin $d$ derivate.

În ciuda formei continue a legii de variaţie (10.1), trebuie subliniat, pentru mai multă exactitate, că funcţia nu are sens decît pe un domeniu discret, alcătuit din mulţimea valorilor întregi (şi, evident, pozitive) ale argumentului, căci nu poate fi vorba de cuvinte care să fi generat un număr fracţionar de derivate.

Trasarea dreptei s-a făcut fără luarea în considerare a numărului de cuvinte neproductive. Cît de mare ar trebui să fie acest număr pentru a satisface ecuaţia se află uşor introducînd $d = 0$ în relaţia (10.1). Se obţine imediat $ln\, n_{lat}(0) = 7{,}475$, adică $n_{lat}(0) = 1763$, valoare ce se abate cu numai 2% faţă de numărul total real de cuvinte

[1] Procedeu matematic de trasare a unei curbe/drepte printr-un nor de puncte experimentale în aşa fel încît suma pătratelor abaterilor dintre valorile măsurate şi cele calculate să fie minimă.



Scanned with OKEN Scanner

de origine latină (1802). Teoretic, ar trebui să existe un număr de $n_{lat}(0) - n_{lat}(1)$ cuvinte neproductive, respectiv 1763 – 1401 = 362, în loc de 474 cîte cuprinde dicţionarul. Decalajul este destul de însemnat dar, după cum se observă, deficitul relativ de termeni avînd cel puţin un derivat (1328 în loc de 1401, cîţi ar rezulta din ecuaţie) compensează în bună parte surplusul de unităţi sterile, aducînd punctul corespunzător abscisei nule foarte aproape de dreapta (10.1) (ordonata 7,50 faţă de 7,475). Vom avea prilejul mai departe să constatăm că lucrurile nu stau la fel şi în cazul cuvintelor de alte origini, ceea ce justifică a posteriori trasarea dreptelor caracteristice numai prin punctele asociate lexicului productiv[1].

Punctul de intersecţie a dreptei din figura 10.1. cu axa orizontală are o semnificaţie deosebită. El corespunde valorii $n_{lat} = 1$, iar abscisa lui reprezintă numărul de derivate pe care ar trebui să îl posede, conform modelului matematic, cel mai productiv cuvînt de origine latină din limba română. Egalînd cu zero membrul drept al ecuaţiei (10.1) obţinem $d = 32{,}5$. Cu alte cuvinte, numărul maxim teoretic de derivate la care ne putem aştepta din partea unei unităţi lexicale aparţinînd fondului latin moştenit este de cca 32-33. De fapt, avînd în vedere că numărul de cuvinte nu poate fi decît întreg, intervalul căutat rezultă ceva mai mare. Orice valoare a $n_{lat}$ cuprinsă între 0,5 şi 1,5 se rotunjeşte, normal, la 1. Logaritmii corespunzători acestor cifre sînt – 0,693 şi 0,405, valori care introduse în ecuaţia (10.1) conduc la o plajă de variaţie a argumentului $d = 31 - 36$. Între aceste limite trebuie să se situeze deci productivitatea maximă a termenilor latini, prognoză confirmată de lista celor mai prolifice cuvinte, unde primele ranguri sînt ocupate de substantivul *faţă* (35 de derivate) şi adjectivul *bun* (32 de derivate).

## 10.2. Cuvinte comune cu albaneza

Conform tabelului 8.3. un număr de 85 de cuvinte-bază au fost apropiate de termeni asemănători (fonetic şi semantic) din limba albaneză. Ar mai fi de adăugat acestora cîteva elemente cu dublă sau

[1] De fapt, din respect pentru adevăr, trebuie spus că, în cursul cercetării, autorul a utilizat iniţial şi punctul corespunzător numărului total de cuvinte, reluînd determinarea ecuaţiilor în momentul în care a descoperit ecarturi flagrante ale acestui punct în raport cu dreptele respective.



Scanned with OKEN Scanner

multiplă etimologie incertă (!) în sensul că posedă în două sau mai multe limbi diferite corespondente semnalate în dicţionare (DLRM şi DEX) sub cf., ca de exemplu: *tărăboi* cf. alb. *tërboz*, ngr. *thórybos; scrum* cf. alb. *shkrump*, bg. *skrum; stavă* cf. alb. *stavë*, bg. *stava; gard* cf. alb. *gardh*, sl. *gradu; baligă* cf. alb. *balgë*, scr. *balega;* şi chiar *na* cf. alb. bg. ngr. magh. *na, nană* cf. alb. *nanë*, bg. scr. *nana*. În total ar fi vorba deci de 92 de cuvinte, dintre care 62 (67,4%, adică mai puţin decît pentru fondul latin) sînt productive. Iată lista acestora:

| | | |
|---|---|---|
| 1. | moş | 38 de derivate |
| 2. | fărîmă | 21 de derivate |
| 3 | groapă | 15 derivate |
| 4. | copil | 14 derivate |
| 5. | murg | 13 derivate |
| 6-7. | brad, gata | 12 derivate |
| 8 | cioc | 10 derivate |
| 9-13. | curma, fluier, fluture, mînz, ţarc | 9 derivate |
| 14-15. | mugure, văpaie | 8 derivate |
| 16-19. | căciulă, măgar, raţă, strungă | 7 derivate |
| 20-24. | bucura, bulz, gălbează, grumaz, grunz | 6 derivate |
| 25-32. | abur, brîu, buză, gard, ghimpe, grapă, vatră, zgură | 5 derivate |
| 33-35. | olog, scrum, traistă | 4 derivate |
| 36-39. | mazăre, pălărie, pîrău, zgardă | 3 derivate |
| 40-51. | baligă, căpuşă, cioară, copac, hămesi, măgură, mire, năpîrcă, pală, sîmbure, şopîrlă, ţap | 2 derivate |
| 52-62. | balaur, barză, ciut, coacăză, mal, mărar, mînzare, păstaie, rînză, stavă, viezure | 1 derivat |

Din pricina numărului total relativ mic de cuvinte, volumul claselor de egală productivitate nu mai creşte monoton cu rangul, deşi în medie tendinţa se menţine. Sistematizate în tabelul 10.2. (analog cu 10.1.)

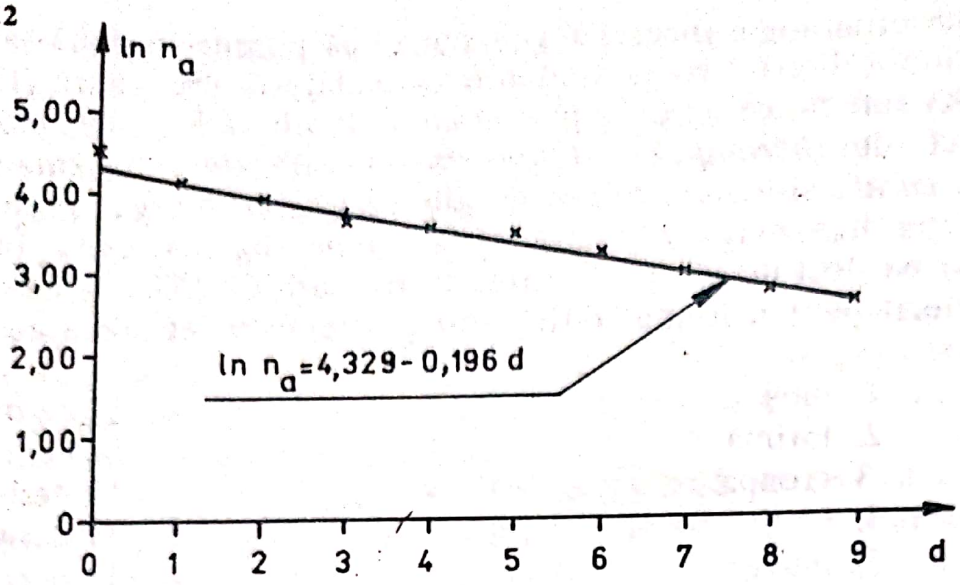
*Tabelul 10.2*

| Nr. derivate ≥ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 92 | 62 | 51 | 39 | 35 | 32 | 24 | 19 | 15 | 13 |
| ln (Nr. cuv.-bază) | 4,52 | 4,13 | 3,93 | 3,66 | 3,56 | 3,47 | 3,18 | 2,94 | 2,71 | 2,56 |



Scanned with OKEN Scanner

FIG.10.2




şi reprezentate grafic în figura 10.2., punctele corespunzătoare celor 9 cercuri succesive care posedă cel puţin 10 elemente se dispun şi ele de-a lungul unei drepte a cărei ecuaţie rezultă:

$$ln\, n_a = 4{,}329 - 0{,}196\, d \qquad (10.2)$$

Procedînd ca în paragraful 10.1., găsim că, potrivit dreptei caracteristice, numărul de cuvinte neproductive ar trebui să fie $ln\, n_a(0) - ln\, n_a(1) = 76 - 62 = 14$, mai puţin de jumătate din cîte există în realitate (30). Trebuie să ţinem însă seama de faptul că elementele comune cu albaneza nu ne dau decît o imagine parţială a influenţei substratului asupra vocabularului limbii române. Cu prilejul examinării situaţiei cuvintelor de origine necunoscută vom relua problema, încercînd să aducem unele corecţii bazate pe analiza comparativă a ecuaţiilor stabilite, astfel încît să includem în calcule şi eventualele relicte lexicale preromane care nu au lăsat urme în albaneză.

### 10.3. Elementul vechi slav

Cuprinde conform tabelului 8.3. un total de 1101 cuvinte nederivate, dintre care 676 sînt productive (61,5% — mai puţine decît în cazul componentelor primordiale ale vocabularului):

| | |
|---|---|
| 1. prost | 20 derivate |
| 2-3. gol, ţintă | 19 derivate |



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 4. | potrivă | 18 derivate |
| 5-6. | lipi, ţigan | 17 derivate |
| 7-9. | drag, pod, suci | 16 derivate |
| 10-12. | clipi, tîrî, val | 15 derivate |
| 13. | rob | 14 derivate |
| 14-17. | drum, sfînt, sticlă, trăsni | 13 derivate |
| 18-21. | obraz, pestriţ, prund, topi | 12 derivate |
| 22-26. | boier, pîndi, pleată, plug, prieten | 11 derivate |
| 27-37. | babă, băl, clăti, mîndru, mladă, noroc, rumen, sfat, tîrn, vesel, vîrf | 10 derivate |
| 38-56. | bold, brazdă, buche, ciocan, cîrciumă, colac, crai, croi, cumpănă, învîrti, nărav, opri, pofti, poveste, săvîrşi, taină, troian, trup, vifor | 9 derivate |
| 57-77. | bici, citi, cîlţi, clei, coajă, gît, glas, milă, ocol, păianjăn, popă, praf, sfii, slab, spoi, struji, ungur, var, veste, vîlvă, zgîrci | 8 derivate |
| 78-102. | cleşte, daltă, dijmă, dobitoc, dragoste, iubi, înveli, metoc, munci, nădejde, năvăli, nevastă, pizmă, pîslă, plesni, potcă, prii, rînd, strună, suliţă, trăi, verigă, vinovat, vrăji, zdrobi | 7 derivate |
| 103-149. | bogat, călugăr, clocoti, dîrz, fală, grămadă, iscodi, istov, îmblăti, încovoia, îngrădi, năuc, neamţ, nevoie, otrăvi, oţel, oţet, păzi, pecete, pipăi, pîrî, pomeni, povară, praştie, prăji, proroc, risipi, schimnic, scîrbă, scump, sfîrşi, spor, stîlp, strung, şopti, ştirb, temei, tocmi, topor, trebui, trîmbiţă, ţeavă, vişină, vîslă, vrăjmaş, zidi, zori | 6 derivate |
| 150-214. | basm, băjenie, camătă, capie, cerni, cinste, cîrmă, cîrpi, clădi, cojoc, colind, corabie, cosi, cumătră, diavol, goni, gunoi, iute, jindui, lacom, milostiv, mită, mînji, movilă, năzări, obişnui, obşte, orîndui, pilă, pîclă, pîrgă, plămădi, pogorî, polei, popri, pribeag, primejdie, privi, rac, rod, sădi, scînci, scrînti, sfinţi, sminti, smîrcîi, stih, strajă, tălmăci, temniţă, tină, tîngui, toci, ţeapăn, ulei, uncrop, viteaz, vîrstă, | |



Scanned with OKEN Scanner

vornic, vreme, zare, zăvor, zdravăn, zîmbi, zmeu — 5 derivate

215-304. armean, bîrnă, bulgar, ciudă, cîrn, clopot, coş, ctitor, danie, dărui, deal, deosebi, dobîndi, doborî, dospi, dumbravă, gloată, greşi, grindă, groază, hîrtie, hohoti, ispăşi, ispiti, isteţ, izbi, îndrăzneală, jeli, jelui, jgheab, jintui, leac, lene, lopată, lovi, milui, mirui, molcom, năduf, năpastă, năsădi, născoci, obrinti, oglindi, osteni, ostrov, pagubă, pinten, plăvan, podoabă, poliţă, porni, postav, prăpastie, prigoni, primeni, primi, protopop, răspîndi, rîcîi, robotă, sili, silnic, sită, sluji, smuci, smochină, snop, soroc, sprijin, staroste, stavilă, stîlci, stînjeni, straşnic, strădui, teşi, tihni, tîmp, tîmpină, tîrg, uliţă, vadră, veşnic, vină, vlagă, volnic, vrabie, vrednic, zadar — 4 derivate

305-411. agheasmă, căli, ceas, cinsti, ciolan, cîrmi, cleveti, cositor, coteţ, dar (subst.), diac, dojeni, dovedi, dungă, gîngav, harnic, hirotonie, hirotonisi, hrăni, huli, iscusi, isprăvi, izbăvi, izbîndi, jar, lebădă, lenevi, luncă, maşteră, măguli, mărturisi, măslină, meni, mirean, mireasmă, molcom, mîrşav, mocni, monah, mreajă, muncă, năimi, năprasnic, nătîng, năvod, nerod, oblu, obloji, ocărî, ocroti, odraslă, osebi, osîndi, păgubi, păli, piron, pîlc, pocăi, podgorie, ponor, poreclă, porunci, postelnic, potopi, pravilă, prescură, prigori, prispă, probozi, puhoi, pustnic, război, răzvrăti, rîvnă, rodi, ropot, scîrnă, sfadă, sihlă, slăvi, slei, slugă, slujbă, slujnică, smoală, smredui, somn, spovedi, stîrc, stolnic, strivi, şopot, tablă, teleagă, tipic, trestie, trudă, voie, voievod, vrajbă, vreasc, vremelnic, zăbovi, zălog, zid, zidar, zvon — 3 derivate

412-534. afurisi, apostol, beli, beznă, blagoslovi, bob, boli, brad, brodi, brutar, canon,



Scanned with OKEN Scanner

caznă, căi, căina, chelar, chică, chilie, cobe, colivă, cocoş, cotei, cotoc, crîng, datină, dreavă, drojdie, duhovnic, evreu, făli, ghizd, gloabă, greabăn, grohoti, hărăzi, ibovnic, icoană, idol, ivi, izmeni, izvodi, izvor, jalbă, jale, jimbat, joardă, libov, logodi, migdal, mirosi, mîrţină, mlaştină, nădrag, năpădi, năpusti, năsli, năvădi, nemernic, obadă, obidi, obijdui, oblădui, ogradă, omorî, ostoi, pistrui, pivniţă, plată, plavie, plăti, pleavă, ploconi, poală, pomană, ponosi, posmag, pospăi, poticni, potrivnic, prag, price, pripas, prohodi, propovădui, razna, răcni, răni, răstigni, răzbi, războinic, rogoz, roi, rufă, schit, scopi, scovardă, sfert, smead, smeri, sol, stareţ, streaşină, stricni, şumăn, tabără, tartor, tescui, ticsi, tîlcui, tînji, tîrgui, treabă, trezi, trîmbă, trîntor, troscot, ucenic, undiţă, veac, vodă, vraci, zăbrea, zămisli, zăticni — 2 derivate

534-672. agud, aievea, anafură, arhiereu, becisnic, belciug, bivol, blajin, blid, bolniţă, boştină, breaslă, buiac, cădelniţă, cădi, ceată, chinovie, cireadă, cislă, ciudi, clucer, coasă, colibă, cosor, covaci, dajdie, destoinic, dezgovi, diacon, dihanie, dîră, dosădi, duh, găvozdi, gîtlan, gleznă, glumi, grajd, grind, hobot, hrean, iaz, icre, iezer, izgoni, jertfă, jitniţă, lăstun, leat, lighioană, liturghie, logodnă, mătanie, medelniţă, mitropolit, mraniţă, mucenic, netrebnic, nevolnic, năsîlnic, novac, obidă, obşti, ogheal, omăt, orez, ovăz, pacoste, peşteră, piscui, pistru, plean, plocon, poftori, poiană, poiată, pojar, poligni, pomăzui, popas, porfiră, posădi, potir, povarnă, pravoslavnic, prăznui, priboli, pricăji, pridvor, pristav, pristăvi, prizări, proslăvi, răstoci, răzbici, războli, rîs (zool.), rît, sad, satană, scoică, scorpie,



Scanned with OKEN Scanner

scund, serafim, sfeclă, sfeşnic, silui, smîrc, smîrd, snoavă, sobor, sobornic, sodom, steblă, stoci, stog, strădanie, sucnă, şiştar, teslar, tocmai, toiag, trainic, trişte, trudnic, ţivli, ţîrcovnic, undi, urni, vădi, vel, vîrî, vraf, zavistnic, zavistie, zavistui, zglobiu, zloată 1 derivat

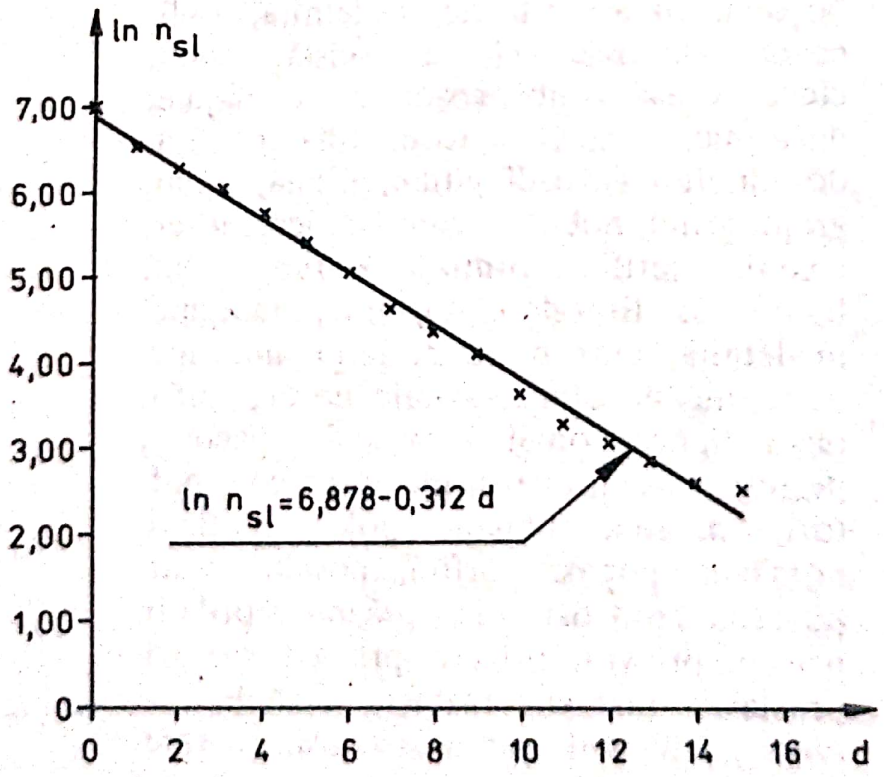
Adăugarea celor cîteva unităţi cu etimologie mixtă paleo- şi neoslavă (*proţap* < *pročepu*, bg. *procep* şi celelalte menţionate la pag. 59) nu modifică semnificativ acest tablou. Ţinînd totuşi seama şi de ele, situaţia generală a elementului vechi slav este cea sintetizată în tabelul 10.3.1. şi ilustrată de figura 10.3.1, unde ne întîlnim din nou cu o distribuţie acuzat liniară.

*Tabelul 10.3.1.*

| NR. DERIVATE ≥ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 1 110 | 676 | 535 | 412 | 305 | 214 | 149 | 102 |
| ln (Nr. cuv.-bază) | 7,01 | 6,52 | 6,28 | 6,02 | 5,72 | 5,37 | 5,00 | 4,62 |

| NR. DERIVATE ≥ | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 77 | 56 | 37 | 26 | 21 | 17 | 13 | 12 |
| ln (Nr. cuv.-bază) | 4,34 | 4,03 | 3,61 | 3,26 | 3,04 | 2,83 | 2,56 | 2,48 |

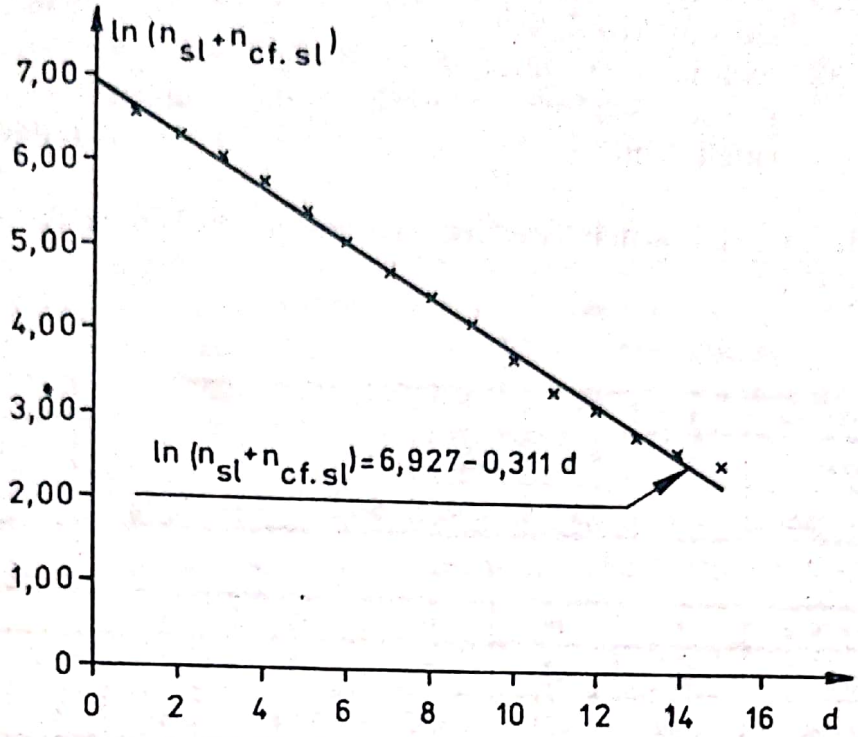
**FIG.10.3.1.**






Scanned with OKEN Scanner

FIG. 10.3.2.




Metoda celor mai mici pătrate furnizează ecuaţia:

$$ln\, n_{slav} = 6.878 - 0{,}312\, d \qquad (10.3.1)$$

Urmărind felul cum se distribuie punctele „experimentale" de-a lungul dreptei de mai sus (v. fig. 10.3.1.), este de semnalat o anumită „depresiune" în zona productivităţilor 7÷12, ca şi cum din inventarul nostru ar lipsi unele cuvinte de origine slavă. Este foarte posibil să şi fie aşa, căci, pe lîngă unităţile cu etimon paleoslav cert, mai există, să nu uităm, şi o serie de termeni cărora dicţionarele le indică proveniența uzînd de prudentul *confer*. Să-i enumerăm pe cei productivi:

| | | |
|---|---|---|
| 1. | pit(ic) | 14 derivate |
| 2. | tăvăli | 12 derivate |
| 3—4. | flăcău, ţeapă | 10 deriyate |
| 5—7. | propti, stropi, zăpadă | 8 derivate |
| 8—9. | perie, preajmă | 7 derivate |
| 10. | zdreanţă | 6 derivate |
| 11. | găsi | 5 derivate |
| 12—14. | necaz, steag, vînj | 4 derivate |
| 15—17. | iscăli, opaiţ, părăsi | 3 derivate |



Scanned with OKEN Scanner

18—25. crîmpei, hudă, înfiripa, reavăn, semeţ, sloveni, trînd, votru — 2 derivate

26—38. bălărie, biet, gătej, jirav, minge, obleagă, prosie, sărăntoc, sipică, şular, ujujit, valah, vinicer — 1 derivat

Cumularea celor două liste are ca rezultat mulţimea de valori din tabelul 10.3.2.

*Tabelul 10.3.2.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 714 | 560 | 429 | 319 | 225 | 159 | 111 | 84 |
| ln (Nr. cuv.-bază) | 6,57 | 6,33 | 6,06 | 5,77 | 5,42 | 5,07 | 4,71 | 4,43 |

| NR. DERIVATE ≥ | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 60 | 41 | 28 | 23 | 17 | 14 | 12 |
| ln (Nr. cuv.-bază) | 4,09 | 3,71 | 3,33 | 3,14 | 2,83 | 2,64 | 2,48 |

În reprezentarea grafică a acestor cifre (v. fig. 10.3.2)este de reţinut că punctele de abscise 7, 8 şi 9 au ajuns practic pe dreapta caracteristică, a cărei ecuaţie a devenit acum:

$$ln\, n_{slav} = 6{,}927 - 0{,}311\, d \qquad (10.3.2)$$

Faptul reprezintă, dacă nu o dovadă, cel puţin un indiciu că o parte însemnată a cuvintelor din lista a doua sînt într-adevăr de origine slavă veche (cu un plus de certitudine pentru termenii care au dat 7—12 derivate pe teren românesc). Singurul punct ce se mai abate sesizabil de la poziţia teoretică rămîne cel corespunzător cuvintelor cu 11 derivate. Am putea continua, fireşte, pe calea sugerată, căutînd ce alte cuvinte româneşti de productivitate 11 sau ceva mai mare pot fi de obîrşie slavă veche (un candidat posibil este *ţelină,* considerat ca provenind din *bg.scr. celina* şi care are în DLRM exact 11 derivate (!): *desţeleni, desţelenire, înţeleni, înţelenire, înţelenit, înţelenitor, înţelina, înţelinare, ţelinioară, ţeliniş, ţelinos*). Să nu cădem însă în greşeala de a cere unei metode prin esenţa ei probabilistă şi globală să ne furnizeze rezultate deterministe, individuale. Puterea explicativă şi relevanţa modelelor discutate în lucrarea de faţă se opresc în faţa specimenului, a cazului particular, întrucît acesta se poate supune legilor speciei lui, după cum la fel de bine se poate constitui în excepţie care să confirme



Scanned with OKEN Scanner

prin abatere comportarea regulată a congenerilor săi. De altfel, obiectivul nostru de căpetenie este tocmai abordarea holistică, globală a unor aspecte ce au avut parte pînă astăzi aproape exclusiv de o tratare atomistică (fără rezultatele căreia, trebuie să o spunem însă, direcţia propusă aici ar fi rămas un simplu deziderat).

Se mai cuvine remarcat că, deşi numărul de cuvinte a crescut odată cu luarea în considerare a unităţilor de origine paleoslavă îndoielnică din a doua listă, panta dreptei a rămas practic neschimbată (tangenta trigonometrică a unghiului pe care ea îl face cu axa absciselor este acum $m_2 = -0{,}311$ faţă de $m_1 = -0{,}312$ cît era iniţial !). Ceva mai departe, cînd vom analiza în detaliu dependenţa pantei de vechimea materialului lexical considerat, vom înţelege şi semnificaţia acestei invarianţe: ea indică (anticipăm) că, în medie, cuvintele cuprinse în cele două liste sînt de aceeaşi vîrstă şi că, prin urmare, avem de a face şi în al doilea caz tot cu termeni aparţinînd stratului vechi slav.

Întrucît coeficientul argumentului $d$ reprezintă panta dreptei respective, primul termen al diferenţei ce alcătuieşte membrul drept al ecuaţiei corespunde valorii funcţiei pentru $d = 0$, adică este egal cu logaritmul numărului total teoretic de cuvinte de originea considerată. Conform ecuaţiei 10.3.1. ar fi trebuit să găsim în dicţionar $e^{6{,}878} = 971$ de cuvinte de obîrşie slavă veche, în loc de 1110 cîte se află acolo efectiv. Un fenomen asemănător a apărut în cazul substratului. Rămîne să îl urmărim în continuare şi să îi căutăm explicaţia la momentul potrivit.

### 10.4. Cuvinte cu etimologie dublă bulgaro-sîrbocroată

Ne ocupăm de ele în imediata prelungire a elementului vechi slav, de care sînt dificil de departajat. După revizuirea pe baza indicaţiilor etimologice din DEX a întregului material lexical neoslav, ierarhia termenilor productivi de sorginte mixtă bulgaro-sîrbocroată a rămas următoarea:

| | | |
|---|---|---|
| 1—.2 | colţ, dascăl | 15 derivate |
| 3. | voinic | 14 derivate |
| 4. | ţelină | 13 derivate |
| 5. | sărac | 10 derivate |



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 6. | potcoavă | 9 derivate |
| 7—10. | gospodar, grădină, puf, ţesală | 8 derivate |
| 11—15. | ciopli, grăi, potlog, prăsi, scoabă | 7 derivate |
| 16—21. | chirie, haiduc, pîrli, pîrpăli, prăvăli, sur | 6 derivate |
| 22—31. | obosi, opări, pleaşcă, prăsi, pritoci, răchită, roi, rujă, scrîşni, tropot | 5 derivate |
| 32—39. | buruiană, dogori, haină, maică, polog, potoli, tărîţe, zaveră, | 4 derivate |
| 40—51. | buşi, cetină, chicot, claie, făclie, ogor, opăci, rasol, samar, scrobi, surlă, zăpuşi | 3 derivate |
| 52—77. | bobot, ceasornic, cobi, covăsi, crap, gaidă, graniţă, jir, jitar, pîrloagă, ploscă, poteră, povîrni, prepeliţă, răboj, răsplăti, sărman, slad, spiţă, stobor, sucală, şipcă, tocilă, trupină, zădărî, zob | 2 derivate |
| 78—116. | babiţă, baştină, boişte, busuioc, ceucă, ciocan (păhărel), cîrpă, cladă, clăbuc, colnic, copaie, drob, gorun, leasă, mătăcină, momi, mreană, obor, opor, otavă, pogonici, răgace, răsădi, răzor, schelă, slănină, sticlete, stropoli, ştiucă, tigaie, tigvă, tîrnăcop, tolbă, tuli, uium, varniţă, vrăbete, zăduf, zlătar | 1 derivat |

Adăugînd cele cîteva unităţi cu etimologie multiplă (de fapt, singurul care are un cuvînt de spus, datorită productivităţii sale ridicate, este *şcoală,* cu 13 derivate în DLRM, considerat de DEX drept provenind din limbile bulgară, sîrbocroată, polonă şi rusă) se ajunge la distribuţia din tabelul 10.4.

*Tabelul 10.4.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 120 | 80 | 53 | 40 | 32 | 22 | 16 | 11 |
| ln (Nr. cuv.-bază) | 4,79 | 4,38 | 3,97 | 3,69 | 3,47 | 3,09 | 2,77 | 2,40 |

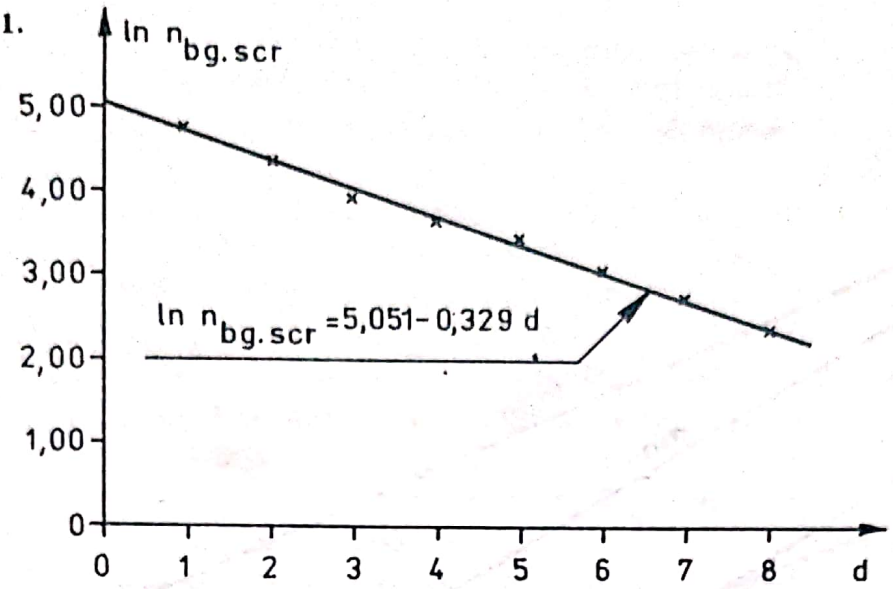
Reprezentate grafic (fig. 10.4.1), punctele ce corespund acestor 8 clase se dispun şi ele aproape perfect coliniar, de-a lungul unei drepte a cărei ecuaţie este:

$$ln\, n_{bg.scr.} = 5{,}051 - 0{,}329\, d \qquad (10.4)$$



Scanned with OKEN Scanner

Fig. 10.4.1.




Teoretic, cel mai productiv cuvînt cu etimologie dublă bulgaro-sîrbocroată ar trebui să aibă în jur de $d = \frac{5,051}{0,329} = 15,4$ derivate, fapt confirmat de prezenţa în capul listei a substantivelor *colţ* şi *dascăl,* ambele de productivitate 15. Tot potrivit ecuaţiei (10.4) numărul total de cuvinte neproductive ar trebui să fie $n_{bg.scr.}(0) - n_{bg.scr.}(1) = 156 - 112 = 44$. În realitate, întîlnim însă 60, din nou mai multe decît numărul scontat, totalul unităţilor lexicale de origine bulgaro-sîrbocroată fiind 180, în loc de 156 cîte ne indică modelul.

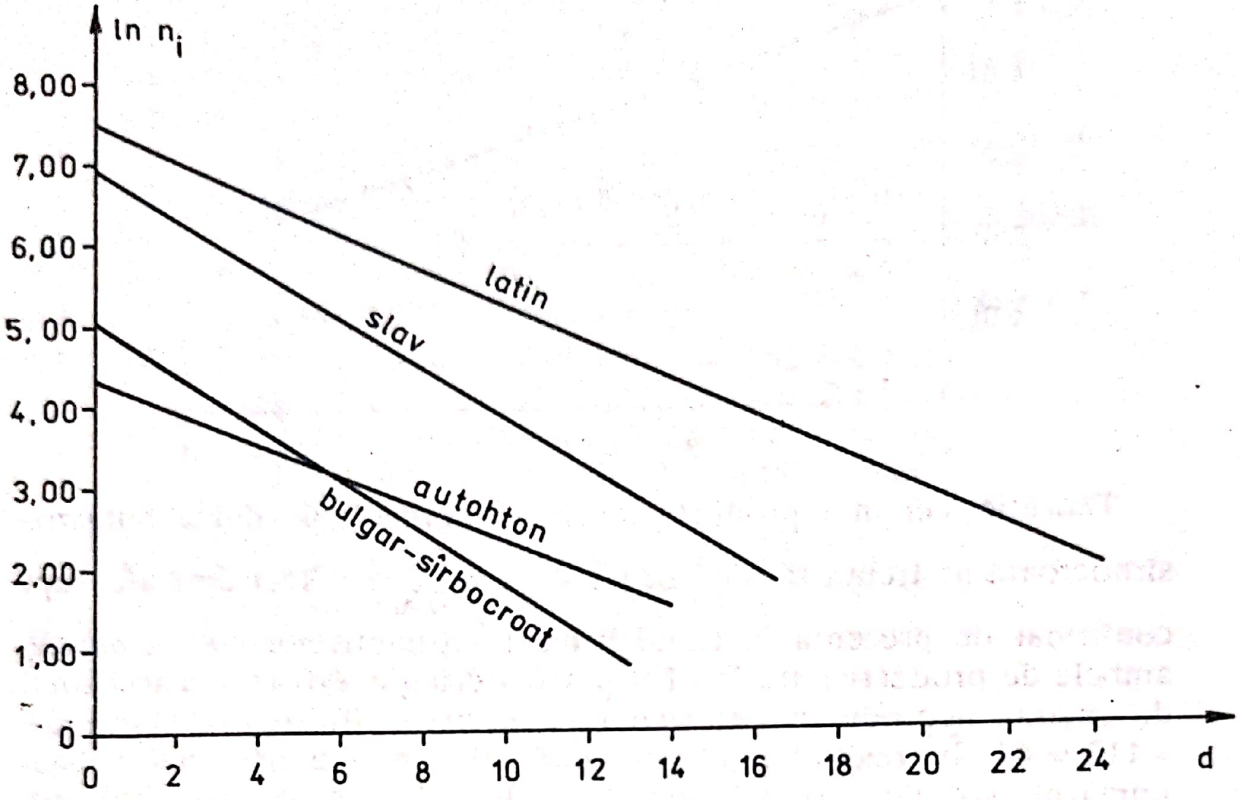
Dacă reprezentăm pe acelaşi grafic cele 4 drepte caracteristice determinate pînă în momentul de faţă (fig. 10.4.2) constatăm că ele sînt două cîte două (aproximativ) paralele. Deşi radical deosebite ca pondere în vocabularul limbii romîne, fondul latin moştenit şi substratul conduc la drepte cu pante apropiate (în jur de 0,2). La fel de pregnantă este similitudinea dintre bogatul element vechi slav şi mult mai restrînsul material lexical cu etimologie mixtă bulgaro-sîrbocroată, reprezentate printr-o pereche de paralele, avînd de data aceasta panta ceva mai mare (puţin peste 0,3). Presupunerea că acest din urmă fond cuprinde încă multe elemente a căror pătrundere în limba română este de dată veche slavă (termeni ca *grădină* sau *opări,* comune cu dialectele sud-dunărene, aparţinînd în mod cert acestei categorii) pare să-şi găsească confirmarea în paralelismul constatat[1].

[1] După cum am mai arătat, D. Gămulescu apreciază, în legătură cu termenii cu dublă etimologie bulgară şi sîrbocroată, că „Marea majoritate sînt vechi sud-slavi" (v. nota noastră 2 de la pag. 59).



Scanned with OKEN Scanner

Fig. 10.4.2




## 10.5. Împrumuturile bulgăreşti

Cuvintele cu etimon bulgar cert (şi unic) din corpusul nostru sînt, în ordinea descrescătoare a productivităţii, următoarele:

| | | |
|---|---|---|
| 1. | grijă | 18 derivate |
| 2. | plută | 14 derivate |
| 3. | crac | 13 derivate |
| 4. | nisip | 12 derivate |
| 5—6. | jupui, lanţ | 10 derivate |
| 7—8. | pită, slugar(nic) | 9 derivate |
| 9—11. | blană, cloci, sfredel | 8 derivate |
| 12—16. | creangă, noroi, pustie, şagă, vopsi | 7 derivate |
| 17—30. | clacă, covrig, gîscă, lăstar, lehamite, meliţă, odihni, plivi, pricină, rînji, sabie, scorni, slobod, suman | 6 derivate |
| 31—44. | arap, bolovan, gîdila, grabă, jilav, lele, lesne, miji, oţări, prăpădi, rudă, sanie, stînjen, vîrtej | 5 derivate |



Scanned with OKEN Scanner

45—62. bolnav, bujor, curcă, desluşi, dumeri, gîlceavă, greblă, grozav, horă, icni, laviţă, nadă, nămete, păstra, pelin, slăbănog, stărui, zălud — 4 derivate

63—85. bălăci, cîrpaci, clonţ, copcă („gaură în gheaţă"), cremene, drişcă, griji, humă, juli, mirişte, năclăi, necăji, nimeri, opincă, păstrăv, pogon, potecă, precupeţ, prilej, primiti, rariţă, rîşniţă, sineală, sorcovi — 3 derivate

86—103. baniţă, borcan, breaz, călină, căpiţă, copcă (vestim.), cosiţă, dropie, gîrbov, morcov, năpîrli, papară, răzgîia, sfoiag, streche, suveică, şiştav, tîrgoveţ — 2 derivate

104—150. baboi, bîtlan, castravete, ciubăr, ciudat, clenci, clisă, clondir, cloşcă, coropişniţă, cosac, crosnă, cută, dîrvar, doică, gîrlă, gligan, godînac, griv, ierarh, izbucni, japcă, jeg, leucă, mezdri, mirodie, morun, năpîrstoc, obială, padină, pestelcă, plătică, pop, popîlni, praz, prăştină, proclet, raclă, răzor, rogojină, roşcov, scoruşă, sobol, tîrcol, ulei (ştiubei), vrej, zăpor — 1 derivat

Alte 137 de cuvinte de origine bulgărească nu au derivate. Prin urmare raportul dintre fondul productiv şi zestrea lexicală totală este în acest caz de numai 52,26%, cel mai mic întîlnit pînă în clipa de faţă.

Pentru a dispune de o imagine exhaustivă a influenţei bulgăreşti, este necesar să cumulăm lista de mai sus cu cea cuprinzînd cuvintele de provenienţă mixtă bulgaro-sîrbocroată din paragraful precedent şi cu materialul lexical cu etimologie dublă sau multiplă care comportă şi o explicaţie prin bulgară. E vorba de încă 22 de cuvinte, descinse din cupluri de etimoane destul de variate: bulgăresc şi vechi slav *(proţap)*, bulgăresc şi neogrecesc *(naramz)*, bulgăresc şi ucrainian *(mălină)*, bulgăresc şi turcesc *(sinie)*, bulgăresc şi rusesc *(ispravnic)*, bulgăresc şi german *(cană)*, bulgăresc şi polonez *(contăş)*, bulgăresc şi maghiar *(podmol)* etc. Bilanţul final este cuprins în tabelul 10.5.

*Tabelul 10.5.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv-bază | 284 | 192 | 144 | 106 | 78 | 53 | 32 | 22 | 15 | 12 |
| ln (Nr. cuv.-bază) | 5,65 | 5,26 | 4,97 | 4,66 | 4,36 | 3,97 | 3,47 | 3,09 | 2,71 | 2,48 |



Scanned with OKEN Scanner

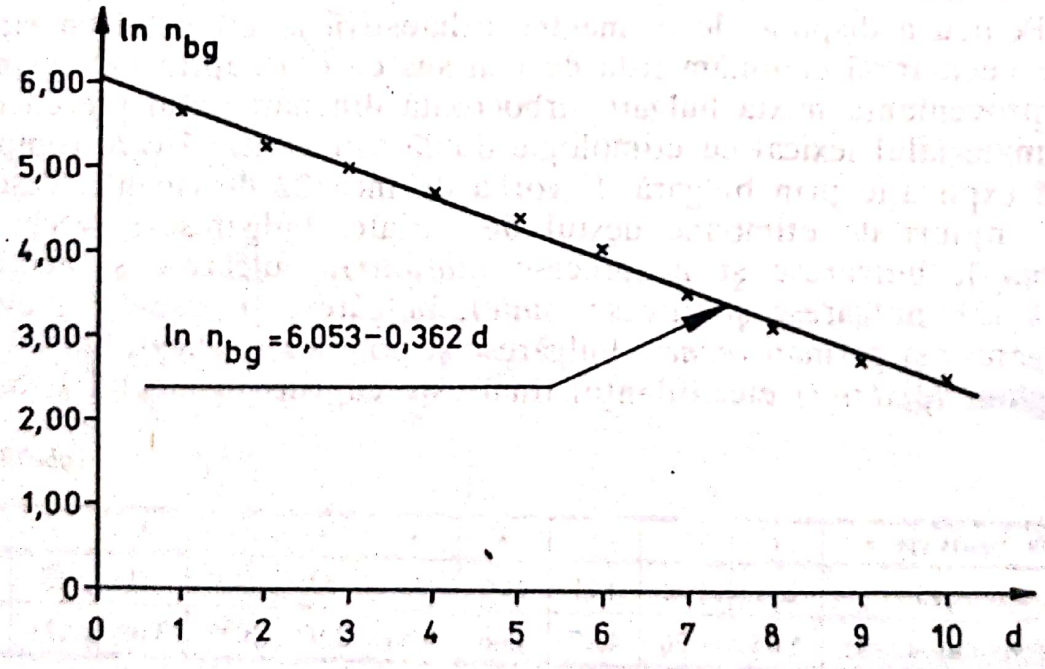
Transpuse grafic (fig.10.5), cele 10 puncte îmbracă mulţumitor dreapta de ecuaţie:

$$\ln n_{bg} = 6{,}053 - 0{,}362\, d \qquad (10.5)$$

Situarea punctelor de abscise $5 \div 6$ puţin deasupra dreptei şi a următoarelor 3 puncte dedesubtul acesteia poate sugera prezenţa în zona productivităţilor $5 \div 6$ a unor cuvinte prebulgăreşti (vechi slave) ca şi existenţa unor termeni mai productivi de origine bulgară rămaşi în afara listelor considerate (ne gîndim în primul rînd la cuvinte cu etimologie incertă ca *broboadă* (cf. bg. *podbradka*), *bumbac (cf. bg. bubak, scr. bumbak), zimţ (cf.bg. zăbec, scr. zubac)*, toate posesoare în DLRM a cîte 8 derivate). Oricum, nerespectarea întrutotul a dependenţei liniare constituie întotdeauna un indiciu că materialul lexical utilizat e întrucîtva eteroclit.

Conform ecuaţiei (10.5), cel mai productiv împrumut din limba bulgară ar trebui să aibă puţin mai mult de 17 derivate (6,053:0,362 = = 17,2). Condiţia e într-adevăr îndeplinită de substantivul *grijă*. Caracterul mai abrupt al dreptei caracteristice (panta $m = 0{,}362$) decît în cazul aportului vechi slav pare să confirme relaţia inversă cu vechimea pe care ne-a sugerat-o examinarea comparativă a dreptelor determinate anterior.

FIG. 10.5






Scanned with OKEN Scanner

## 10.6. Împrumuturile sîrbocroate

Elementul sîrbesc are o pondere sensibil inferioară celui bulgăresc, cuvintele productive ce-i aparţin în mod neîndoielnic (etimologie sigură şi unică) fiind în număr de numai 50:

| | | |
|---|---|---|
| 1. | raz | 10 derivate |
| 2—3. | boltă, pernă | 7 derivate |
| 4—5. | sîrb, zbeng | 6 derivate |
| 6. | dospi | 5 derivate |
| 7—9. | prăbuşi, puhav, scai | 4 derivate |
| 10—18. | cîrd, cujbă, izină, ogoi, priboi, şiţă, tîrnosi, vrîstă | 3 derivate |
| 19—30. | birt, chiti 1, chiti 2, cotcă, dăinui, doniţă, drug, drugă, pleoşti, pup, ţiglă, zgomot | 2 derivate |
| 31—50. | duhan, iezi, odor, păstrugă, pilar, plută (bot.), pogodi, pretcă, pui (numism.), răzvedi, sagnă, scoc, sfor (hidro.), slin, sloven, snagă, sovîrf, sud (= judecată), şoacăţ, vîrşă | 1 derivat |

Etimologie multiplă au, pe lîngă unităţile inventariate sub 10.4., un număr de încă 40 de cuvinte, prezentînd combinaţii variate ca: scr. & maghiar (10 cuvinte, printre care *pahar* cu 9 derivate, *şubă* cu 3, *megieş* cu 2), scr. & pol. *(panţir, şliboviţă)*, scr. & ucr. *(laz, şovar)*, scr. & rus *(batog, crug)*, scr. & germ. *(bocal, şteap)*, scr. & franc. *(guzlă)*, scr. & magh. & germ. *(păstîrnac)*, scr. & pol. & ucr. *(poleac)*, scr. & pol. & rus *(silitră)*.

Totalizarea acestor aporturi conduce la cifrele din tabelul 10.6.1. şi la reprezentarea grafică din figura 10.6.1., unde dreapta caracteristică rezultantă are ecuaţia

$$\ln n_{scr} = 5{,}450 - 0{,}349\, d \qquad (10.6.1)$$

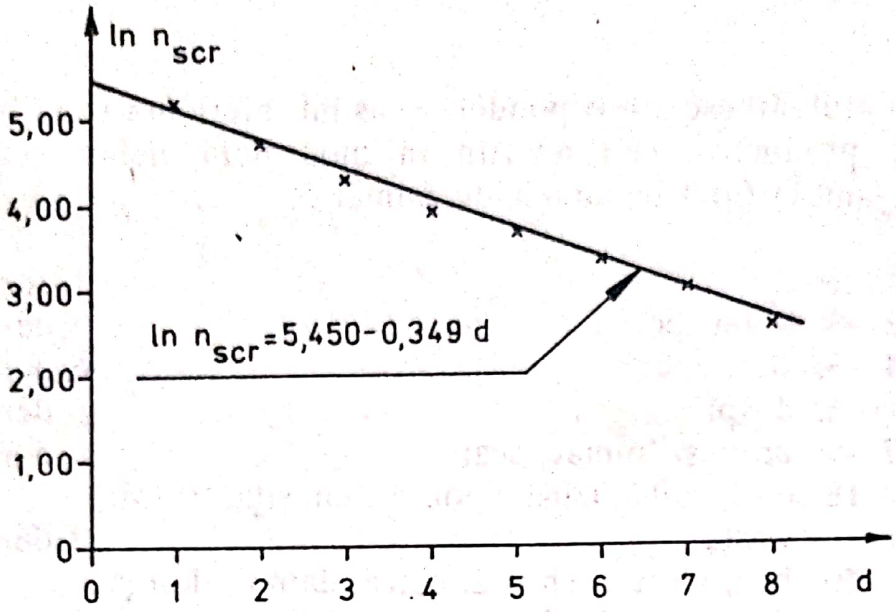
*Tabelul 10.6.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 180 | 113 | 73 | 50 | 39 | 28 | 20 | 13 |
| ln (Nr. cuv.-bază) | 5,19 | 4,73 | 4,29 | 3,91 | 3,66 | 3,33 | 3,00 | 2,56 |



Scanned with OKEN Scanner

FIG. 10.6.1.




Luînd în considerare şi cuvintele lexicalmente sterile, un total de 377 de unităţi au, mai mult sau mai puţin, provenienţă sîrbocroată, aşa încît ponderea în ansamblu a termenilor productivi este de numai 47,7%, inferioară celei ce caracterizează elementul bulgar. De altfel, potrivit ecuaţiei (10.6.1), cuvintele româneşti de origine sîrbocroată ar trebui să fie doar în număr de $e^{5,450} = 233$. Prezenţa în dicţionar a unui număr de unităţi neproductive depăşind cota scontată se explică prin insistenţa DLRM asupra lexicului regional, în care influenţa sîrbească este considerabilă. De altminteri, după sfîrşitul perioadei medievale, influenţa sîrbocroată nu s-a mai exercitat decît asupra graiurilor din Banat şi din zonele strict limitrofe.

O anomalie ce se cere explicată este valoarea mai mică a pantei dreptei (10.6.1) decît cea găsită pentru elementul bulgar, ceea ce contravine succesiunii în timp a celor două influenţe, ştiut fiind faptul că legăturile dintre români şi sîrbi s-au intensificat de-abia după declinul statului bulgar şi că deci, în medie, împrumuturile bulgăreşti sînt cele care ar trebui să rezulte drept mai vechi. Contradicţia e numai aparentă. Ea se datorează faptului că, în absenţa unor cuvinte sîrbeşti foarte productive, ponderea principală în tabelul 10.6.1. revine termenilor cu dublă etimologie bulgaro-sîrbocroată care aparţin în majoritate unei faze mai vechi a contactelor noastre cu slavii sud-dunăreni. Raporturile cronologice normale dintre cele două influenţe se restabilesc dacă facem abstracţie, atît în calculul influenţei sîrbeşti cît şi al celei bulgăreşti, de acest fond comun, analizat de altfel separat în paragraful 10.4.



Scanned with OKEN Scanner

Limitîndu-ne la împrumuturile ce nu prezintă fenomenul dublei etimologii, tabelul şi ecuaţia dreptei caracteristice capătă forma:

*Tabelul 10.6.2.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 164 | 112 | 91 | 66 | 46 | 31 | 16 | 11 |
| ln (Nr. cuv.-bază) | 5,10 | 4,72 | 4,51 | 4,19 | 3,83 | 3,43 | 2,77 | 2,40 |

$$ln\, n_{bg} = 5{,}597 - 0{,}384\, d \qquad (10.6.2)$$

pentru elementul bulgar şi

*Tabelul 10.6.3.*

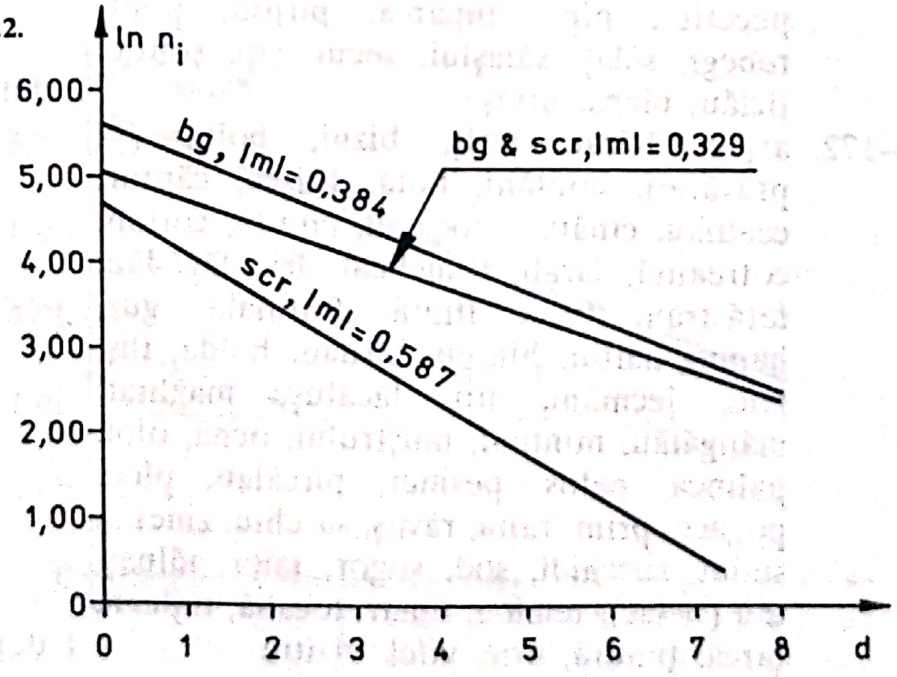
| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nr. cuv.-bază | 60 | 33 | 20 | 10 |
| ln (Nr. cuv.-bază) | 4,09 | 3,50 | 3,00 | 2,30 |

$$ln\, n_{scr} = 4{,}691 - 0{,}587\, d \qquad (10.6.3)$$

pentru cel sîrbocroat.

Dispuse, spre comparaţie, pe aceeaşi diagramă (fig. 10.6.2) dreptele (10.4.1), (10.6.2) şi (10.6.3) au într-adevăr înclinări progresiv mai mari cu cît se referă la influenţe lingvistice mai recente. Valoarea absolută a pantei creşte de la 0,329 pentru fondul lexical comun bulgaro-sîrbocroat, la 0,384 în cazul împrumuturilor bulgăreşti şi la 0,587 în al celor sîrbocroate.

FIG. 10.6.2.






Scanned with OKEN Scanner

## 10.7. Elementul maghiar

Ordonate pe criteriul productivităţii, cuvintele de origine maghiară consemnate în DLRM sînt:

| | | |
|---|---|---|
| 1. | şoim | 13 derivate |
| 2—3. | gînd, socoti | 11 derivate |
| 4—5. | ham (subst.), hotar | 9 derivate |
| 6—13. | cătană, chip, locui, meşter, meşteşug, oraş, seamă, viclean | 8 derivate |
| 14—16. | chibzui, şindrilă, tăgădui | 7 derivate |
| 17—19. | harţă, neam, zăbală | 6 derivate |
| 20—30. | alcătui, altoi, bănui, chin, cizmă, întîlni, lacăt, mîntui, ogar, ponci, vamă | 5 derivate |
| 31—48. | bardă, beteag, birui, bîntui, căpui, cheltui, corci, gazdă, haită, iobag, îngădui, larmă, panglică, pildă, ponoslui, sîrg, uliu, urlui | 4 derivate |
| 49—75. | baie (= mină), belşug, budăi, celui, chezaş, ciopată, ciuf, coroi, cortel, făgădui, fel, grof, husar, jigări, jurui, labă, marfă, mistui, momiţă, nemeş, papistaş, sicriu, tămădui, tobă, ţinc, vameş, vînzoli | 3 derivate |
| 76—105. | bir, bundă, burzului, dîmb, dorobanţ, drot, feşteli, gelui, guler, înverşuna, lăfăi, majă, mohor, nesăbuit, orbalţ, pandur, pecetlui, pipă, piparcă, pîrjoli, pont, rebegi, sălaş, sălăşlui, secui, şip, şontîc, ţiclău, uiagă, uriaş | 2 derivate |
| 106—172. | arţag, bărat, biriş, bizui, boltă (= prăvălie), bontăni, botă, brişcă, căprar, ceahlău, cinătui, ciocălău, ciurdă, cotlon, cotreanţă, dărab, deşucheat, dric, făgădău, ferăstrău, firez, flintă, fuituială, guz, hamiş, hăitui, hinteu, hîrdău, holdă, iliş, jale, jecmăni, jip, lăcătuş, maghiar, măngălău, minteni, muştrului, ocnă, oloi, palincă, paloş, pesmet, pîrcălab, pleu, pogace, prim, raita, răvaş, saschiu, sinelui, sudui, surchidi, şod, şogor, taier, tălpaş, tău (= lac), teleleu, timar, tocană, tujlean, ţarcă, ţundră, uric, vifel, vizitiu | 1 derivat |



Scanned with OKEN Scanner

Ca şi în cazurile precedente, tabelul prezentat mai jos ţine seama în plus de existenţa unor termeni cu etimologie dublă: magh. & scr. (v. paragraful anterior), magh. & tc. *(surduc)*, magh. & ucr. *(taşcă)*, magh. & rus *(potroc)*, magh. & pol. *(mustră)*, magh. & sas *(ţandără)* etc.

*Tabelul 10.7.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 184 | 112 | 79 | 49 | 31 | 20 | 17 | 14 |
| ln (Nr. cuv.-bază) | 5,21 | 4,71 | 4,37 | 3,91 | 3,43 | 3,00 | 2,83 | 2,64 |

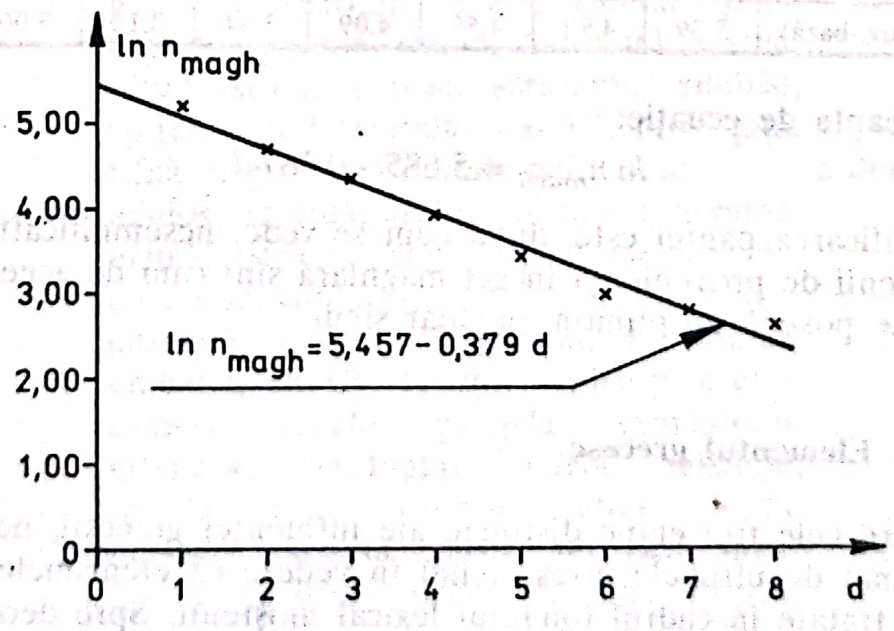
Dreapta construită pe baza acestor date are ecuaţia:

$$ln\, n_{magh} = 5{,}457 - 0{,}379\, d \qquad (10.7.1)$$

Numărul scontat de cuvinte neproductive este $n_{magh}(0) - n_{magh}(1) = 243 - 129 = 105$, de aproape trei ori mai mic decît cel real (292). Rezultatul nu surprinde, dacă observăm că, din totalul de 476 de cuvinte de origine maghiară, 176 sînt regionale sau învechite, marea majoritate a acestora, respectiv 139, fiind cu totul lipsite de derivate. Raportul dintre fondul lexical productiv şi cel total atinge din această cauză o valoare foarte scăzută, de numai 36%.

Cît priveşte panta dreptei din figura 10.7, ea este aproximativ egală cu cea corspunzătoare elementului bulgăresc (considerat fără fondul comun, probabil mai vechi, bulgaro-sîrbocroat): 0,379 faţă de 0,384 (v. ec. (10.6.2)).

**FIG. 10.7**






Scanned with OKEN Scanner

Cum se modifică parametrii dreptei (10.7.1) dacă luăm în considerare şi materialul lexical de origine maghiară incertă? Ne referim la următoarele cuvinte productive:

| | | |
|---|---|---|
| 1. | talpă | 11 derivate |
| 2. | tîlhar | 10 derivate |
| 3. | glod | 7 derivate |
| 4. | sas | 6 derivate |
| 5—6. | balmoş, sîmbră | 5 derivate |
| 7—10. | ademeni, copoi, răzeş, tîrşii | 4 derivate |
| 11—15. | bădăran, bîigui, jumuli, şontorog, tehui | 3 derivate |
| 16—23. | bumb, butaş, cimpoi, ciuciuli, hîrjoni, morocăni, tăpălagă, tărcat | 2 derivate |
| 24—33. | buzna, deschilini, făcălui, ghiob, handralău, hîi, jigodie, mîglă, şomoiog, tupungi | 1 derivat |

şi la cele încă cîteva apropiate deopotrivă de cuvinte maghiare şi nemaghiare (ex.: *cuţu* cf. *bg. scr. kuče, magh. kuszi*, cu derivatele *cuţulache* şi *cuţulan*).

Cumulate cu lista elementelor sigure, acestea conduc la bilanţul din tabelul (10.7.2).

*Tabelul 10.7.2.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 219 | 137 | 95 | 60 | 38 | 24 | 20 | 16 |
| ln (Nr. cuv.-bază) | 5,39 | 4,92 | 4,55 | 4,09 | 3,64 | 3,18 | 3,00 | 2,77 |

şi la dreapta de ecuaţie:

$$ln\, n_{magh} = 5{,}685 - 0{,}387\, d \qquad (10.7.2)$$

Modificarea pantei este, după cum se vede, nesemnificativă, semn că termenii de proveniență incert maghiară sînt cam de aceeaşi vîrstă cu cei ce posedă un etimon maghiar sigur.

## 10.8. Elementul grecesc

Dintre cele trei etape distincte ale influenţei greceşti, ne ocupăm aici numai de ultimele două, avînd în vedere că elenismele latineşti au fost tratate în cadrul fondului lexical moştenit. Spre deosebire de



Scanned with OKEN Scanner

lucrările lexicografice mai vechi, nici DLRM nici DEX nu mai operează distincţia dintre împrumuturile medio- greceşti (bizantine) şi aportul neogrecesc din secolele XVI—XIX. Unităţile productive cărora DLRM le atribuie o ascendenţă neogreacă sînt:

| | | |
|---|---|---|
| 1. | tipar | 12 derivate |
| 2—3. | săpun, sigur | 10 derivate |
| 4—5. | pat, patimă | 9 derivate |
| 6. | smalţ | 8 derivate |
| 7—10. | condei, maimuţă, sfoară, zugrav | 7 derivate |
| 11—14. | bărbier, prisos, stivă, zahăr | 6 derivate |
| 15—20. | argat, cămară, desagă, plapumă, strachină, smeur | 5 derivate |
| 21—40. | aht, alică, argăsi, buzunar, cărturar, chivernisi, cort, dichisi, dram, folos, franzelă, ieftin, lipsi, mărgărit, plictisi, schimonosi, sosi, spătar, trandafir, za | 4 derivate |
| 41—66. | aerisi, agonisi, anost, babacă, boboc, bogasiu, călţun, cărămidă, condică, fandosi, fasole, furtună, lămîie, logofăt, mană, molipsi, oftică, pedepsi, piper, plăsmui, portocală, pricopsi, protipendadă, rodie, sîrmă, tiranic | 3 derivate |
| 67—106. | anatemă, arac, arhondar, arvună, beteală, castan, cofeturi, comis, cucuvaie, egumen, elen, fanariot, fustă, hiritisi, lehuză, marghiol, mărgăritar, navlu, omidă, plastografie, plămîn, salată, saramură, sastisi, sărad, scafă, sclivisi, scrumbie, silabisi, spiţer, splină, stafidă, stambă, tară, ţaţă, ţîfnă, urgisi, varvar, zaharisi, zodie | 2 derivate |
| 107—185. | aghios, aguridă, arghirofil, argos, avramă, azimă, berbant, cais, capelă (= bonetă), catavasie, catehiza, chitru, colţun, costisi, curmală, despot, dragoman, economisi, embatic, enorie, eparhie, epitrop, eterie, evlavie, garafă, garoafă, geamantan, grămătic, hardughie, hartă, iasomie, igrasie, ison, istorisi, lefter, lepră, magazie, masala, misit, mistrie, mosc, mutră, nostim, oleacă, omilie, oropsi, | |



Scanned with OKEN Scanner

ovrei, paner, paraponisi, parohie, pastă, pastramă, păscălie, plic, profir, prohorisi, protoiereu, protoprezbiter, riglă, salţă, saltea, sardea, scufă, scul, sertar, sfeterisi, sihastru, sinchisi, sodomie, stacoj, stol, stradă, taifas, tendă, tiran, trapez, trapezare, tulpan, turlă, zulie — 1 derivat

Faţă de cele 624 de cuvinte „neogreceşti" din dicţionar, fondul productiv reprezintă numai 29,6%, mai puţin chiar decît în cazul elementului maghiar. Cu valorile din tabelul 10.8. se determină ecuaţia dreptei caracteristice reprezentată în figura 10.8., care este:

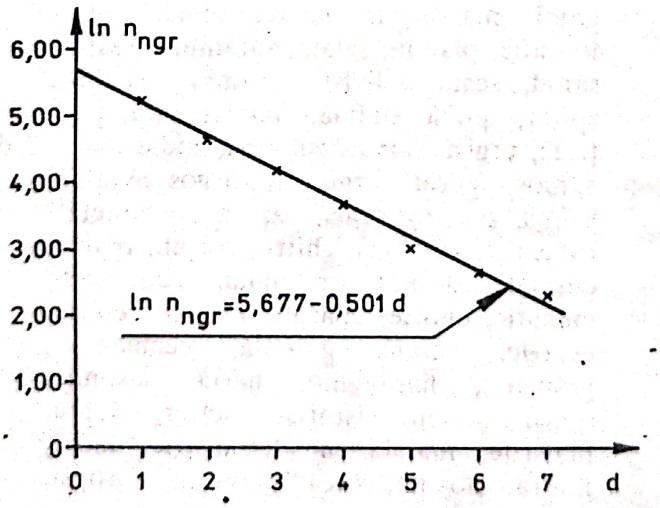
$$ln\ n_{ngr} = 5{,}677 - 0{,}501\ d \qquad (10.8)$$

*Tabelul 10.8.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 185 | 106 | 66 | 40 | 20 | 14 | 10 |
| ln (Nr. cuv.-bază) | 5,22 | 4,66 | 4,19 | 3,69 | 3,00 | 2,64 | 2,30 |

Ca şi în alte situaţii, punctul corespunzător abscisei nule se abate cu mult de la traseul dreptei, din pricina marii discrepanţe existente între numărul scontat şi cel real de cuvinte neproductive. Într-adevăr, ne-am aştepta, potrivit ecuaţiei (10.8), ca numai 153 de cuvinte „neogreceşti" să fie lipsite total de derivate, ceea ce se întîmplă însă cu 439 de cuvinte. Vom vedea repetîndu-se fenomenul pentru toate

FIG.10.8






influenţele l
crescînd cu
este mai mic
fig. 10.1) per
Există o
productive d
două limbi,
cuvinte e rap
redusă. Pe
neproductive
dispar trepta
cîte corespu
caracteristice.
minimă pentr
neologismelo

### 10.9. Elem

Prezintă a
de trei valuri
împrumuturi
turcească ma
XV—XVIII),
şi în cazul ap
furnizează to
turceşti/neogre
conjuncturi p
desuetudine a
Dintre cele
DLRM, următ

1. t
2. c
3—4. t
5—7. c
8—9. c
10—16. c
z

[1] Deşi unele d
etimoanele turceşti



Scanned with OKEN Scanner

influenţele lexicale ce vor mai fi analizate în continuare, decalajul crescînd cu cît e vorba de aporturi mai recente. Faptul că abaterea este mai mică în cazul împrumuturilor mai vechi şi aproape nulă (v. fig. 10.1) pentru fondul latin moştenit s-ar putea interpreta astfel:

Există o proporţie optimă între cuvintele neproductive şi cele productive de o anume origine. În faza iniţială a contactului dintre două limbi, acest raport nu poate fi realizat, deoarece afluxul de cuvinte e rapid (adesea exploziv), iar productivitatea noilor veniţi încă redusă. Pe măsura trecerii timpului, dintre cuvintele iniţial neproductive unele îşi întemeiază o familie, iar altele, rămase izolate, dispar treptat, tendinţa fiind de a nu se conserva decît exact atîtea cîte corespund proporţiei fixe exprimate de ecuaţia dreptei caracteristice. Aşa se face că abaterea punctului de abscisă nulă este minimă pentru cuvintele moştenite şi atinge valori foarte mari în cazul neologismelor.

### 10.9. Elementul turcic

Prezintă analogii cu cel grecesc în sensul că şi în acest caz e vorba de trei valuri succesive: o influenţă cumană (căreia îi aparţin probabil împrumuturi ca *baltag, capcană, cazan, cioban, duşman*[1] etc), una turcească mai veche, de caracter preponderent popular (sec. XV—XVIII), şi alta mai nouă, corespunzătoare epocii fanariote. Ca şi în cazul aportului grecesc, numărul maxim de elemente învechite îl furnizează tocmai stratul cel mai recent, afluenţa de termeni turceşti/neogreceşti din acea perioadă datorîndu-se în principal unei conjuncturi politice, care, odată depăşită, a antrenat căderea în desuetudine a unei părţi însemnate a materialului lexical respectiv.

Dintre cele 1092 de cuvinte nederivate de origine turcească din DLRM, următoarele 259 sînt productive:

| | | |
|---|---|---|
| 1. | turc | 12 derivate |
| 2. | chiabur | 11 derivate |
| 3—4. | tabac (= tăbăcar), tătar | 10 derivate |
| 5—7. | cioban, duşman, ţol | 9 derivate |
| 8—9. | catran, maramă | 6 derivate |
| 10—16. | chef, ciomag, cîntar, darac, găitan, rindea, zeflemea | 5 derivate |

[1] Deşi unele dintre ele sînt atestate în cumană *(cazan, teanc)*, DLRM trimite la etimoanele turceşti, cu singura excepţie a lui *beci* < cuman *bec*



Scanned with OKEN Scanner

17—33. arnăut, boia, bostan, ciubuc, mahala, mazil, oină, para, peltic, perdea, pingea, schingi, suliman, tighel, zapciu, zarzără, zăbun — 4 derivate

34—66. băcan, bursuc, cafea, caftan, casap, călăuză, chelbaş, chior, condur, dulgher, farfurie, fistic, ghidi, hain, lambă, lăută, leş, meremet, olac, palavragiu, papuc, pazvantoglu, rachiu, salahor, serdar, sobă, surghiun, talaz, tutun, tutungiu, vizir, zor, zurba — 3 derivate

67—126. abagiu, amanet, arcan, bulucbaşă, catifea, cazan, ceapraz, cerdac, chel, chenar, chiolhan, chiulangiu, ciob, codoş, conac, cusur, cutie, dambla, dovleac, dugheană, dulap, fudul, haimana, halva, hangiu, harbuz, ienicer, işlic, levent, lichea, mahmur, măscară, mezel, mofluz, moft, murdar, muşama, năstrapă, odaie, ortac, pehlivan, puşti, saiggiu, sidef, soi, sultan, surugiu, şalvari, şerbet, şiret (= viclean), tabla, tarabă, tarla, teanc, telal, toi, tulumbă, uluc, zaraf, zuluf — 2 derivate

127—259. alagea, baltag, basma, băcan (bot.), bărdacă, becher, bina, boccea, boiangiu, bondoc, bucluc, buluc, burghiu, buzdugan, cafegiu, cafenea, caic, caimacam, cange, capac, caraulă, catîr, caval, cazangiu, cazma, cealma, ceauş, chiftea, cherestegiu, chibrit, chihlimbar, chimir, chioşc, ciochină, ciorap, ciortan, ciurec, cîrmîz, cîşlă, coltuc, covată, dandana, dara, derviş, divan, dud, duium, fermenea, fes, fildeş, fotă, gavanos, geaba, geambaş, geantă, ghioci, ghiozdan, giol, giubea, giuvaergiu, gîrbaci, grimea, gugiuman, harabagiu, hareci, haz, hurmuz, huzmet, huzur, iatac, iaurgiu, ibric, ilic, iureş, leafă, lighean, lulea, martac, mazdrac, mămular, micşunea, moftangiu, mosor, moşmoană, mungiu, nazuri, năut, neneacă, nufăr, nur,



Scanned with OKEN Scanner

| | |
|---|---|
| pafta, perciune, peşcheş, peşchir, pilaf, poturi, raft 1 (= poliţă), raft 2 (= harnaşament), ragea, rahat, saca, saftian, sarma, satîr, saxana, seimen, sictir, simigiu, sipet, sufragiu, sunduc, şah, şal, şiret (= şnur), tandur, tarapana, tasma, tavă, taxidar, tembel, temenea, tichie, tinichea, tinichigiu, tipsie, tocmagi, tuci, ursuz, vişinap, zambilă, zarif, zarzavagiu, zevzec | 1 derivat |

Bilanţul din tabelul 10.9. ţine seama şi de cele 5 cuvinte cu etimologie dublă sau multiplă: turcă & bulgară *(samsar, sinie)*, turcă & maghiară *(surduc)*, turcă & neogreacă *(cafas)*, turcă & maghiară & polonă *(palancă)*.

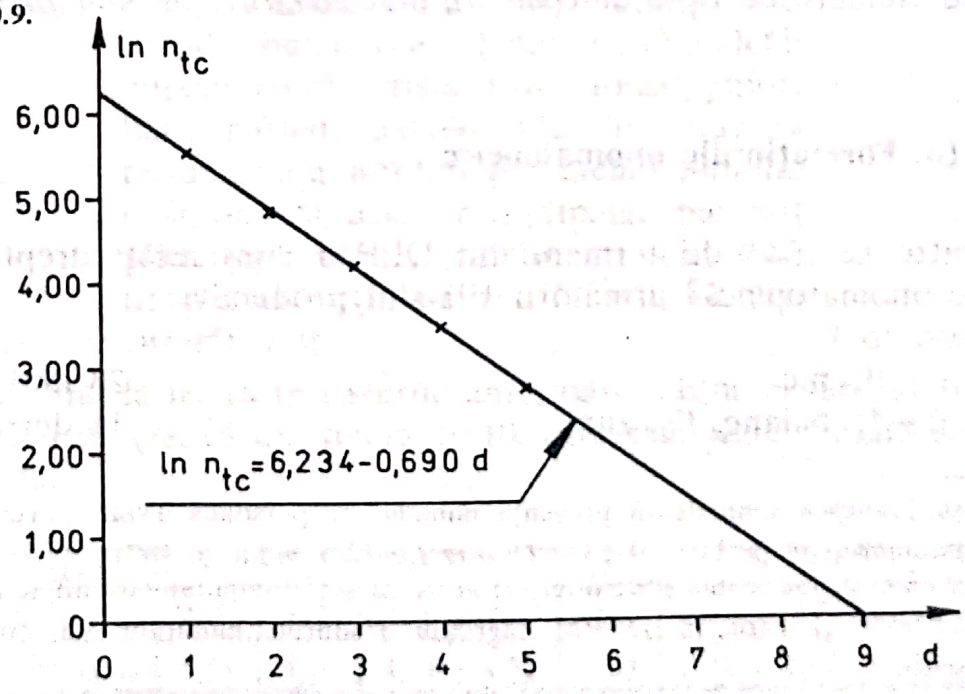
*Tabelul 10.9.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nr. cuv.-bază | 259 | 126 | 66 | 33 | 16 |
| ln (Nr. cuv.-bază) | 5,56 | 4,84 | 4,19 | 3,50 | 2,77 |

Ecuaţia dreptei trasate prin/printre punctele respective (fig. 10.9) are forma:

$$ln\, n_{tc} = 6{,}234 - 0{,}690\, d \qquad (10.9)$$

FIG. 10.9.






Scanned with OKEN Scanner

Conform ecuaţiei, numărul de cuvinte neproductive de origine turcească ar trebui să fie $e^{6,234} - e^{5,544} = 510 - 256 = 254$, adică de mai bine de 3 ori mai mic decît cel efectiv înregistrat, care este 838. Se confirmă astfel din nou observaţiile anterioare privind situaţia specială a punctului corespunzător abscisei nule ca şi faptul că abaterea acestuia de la dreapta caracteristică este cu atît mai mare cu cît se referă la aporturi lexicale mai recente. În contrast cu el, cele 5 puncte corespunzătoare fondului productiv sînt aproape perfect coliniare (v. fig. 10.9). Dreapta pe care ele o determină intersectează axa orizontală în dreptul abscisei $d = 9$; ne aşteptăm deci ca cel mai productiv cuvînt provenit din turceşte să posede 9 derivate. De fapt însă, potrivit listei de mai sus, prognoza este larg depăşită, căci nu doar unul, ci trei cuvinte *(cioban, duşman, ţol)* au cîte 9 derivate, iar alte 4 cuvinte depăşesc şi acest nivel de productivitate: *tabac* şi *tătar* au cîte 10 derivate, *chiabur* — 11, *turc* — 12. Faptul poate fi privit drept o dovadă a apartenenţei la un strat lexical mai vechi cel puţin a cîtorva dintre aceşti termeni, lucru absolut plauzibil, dacă reamintim că pentru *cioban* şi *duşman* s-au propus etimoane cumane, iar etnonimele *tătar*[1] şi *turc* este logic să fi pătruns în românește încă înainte de vremea contactelor lingvistice directe cu turcii. În ceea ce îl priveşte pe *chiabur*, înclinăm să credem că prezenţa în DLRM a unui număr sporit de derivate ale acestuia constituie o supralicitare conjuncturală legată de contextul social-politic în care apărea dicţionarul, moment caracterizat de vehicularea intensă în presă şi la radio a unor alcătuiri lexicale efemere de tipul *deschiaburi, deschiaburire, închiaburit* etc.

### 10.10. Formaţiunile onomatopeice

Printre cei 343 de termeni din DLRM consideraţi drept avînd origine onomatopeică[2] următorii 191 sînt productivi :

| | | |
|---|---|---|
| 1. | pic | 23 derivate |
| 2—4 | balang, fîş, zur | 11 derivate |

[1] Gh. Ivănescu semnalează prezenţa numelui de persoană Tatar, purtat de un boier maramureşean pe la anul 1350 (*Istoria Limbii române* p. 437).

[2] Inventarul ţine seama şi de interjecţiile cu caracter neimitativ asimilate de DEX onomatopeelor şi care în DLRM îngroşau rîndurile unităţilor cu etimologie necunoscută.



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 5—8. | buf, dura, pis, vîj | 10 derivate |
| 9—12. | clamp, puf, scîrţ, zvîc | 9 derivate |
| 13—16. | bîz, hîr, ţîr, vui | 8 derivate |
| 17—19. | hîrşti, sfîr, zbîrn | 7 derivate |
| 20—36. | chiu, chiţ, cirip, cîr, cîrc, clonc, cronc, fleaşc, mîr, mîţă, mor, poc, sfor, şu, ţop, zăngăni | 6 derivate |
| 37—51. | bubui, fîl, fîs, foşni, gîlgîi, hodorog, hurui, miorlăi, pîlpîi, pîr, răpăi, sîsîi, şfichi, ţac, ţiu | 5 derivate |
| 52—74. | behehe, bîjbîi, boc, bufnă, chelălăi, chiorăi, cotcodac, fîţîi, ham, hor, horcăi, hurduca, lipa, mac, miau, of, piu, pleasc, tranc, tronc, ţîşti, zdranc, zgîlţîi | 4 derivate |
| 75—104. | bolborosi, bombăni, ciorobor, clanc, clap, cucurigu, dang, dîrdîi, dup, forfoti, fornăi, gîfîi, gîgîi, grohăi, guiţ, hăi, hîrşcîi, hop, leorbă, molfăi, morfoli, orăcăi, ronţăi, şopa, tio, tropăi, ţist, ţîrlîi, ţoc, zup | 3 derivate |
| 105—140. | bleaşc, caţa, cîrcîi, cling, coviţ, dă, dudui, fonf, gînguri, hap, hai, haţ, hăui, hăuli, hîrcîi, hîţ, horp, hui, iu, jnap, lepşi, miorc, pîc, pleosc, st, şoşoşo, tic-tac, toc, turui, ţah, ţuşti, ugui, zăpăi, zdronc, zdup, zuzui | 2 derivate |
| 141—191. | bang, bălălăi, bîltîc, bîr, bondar, boscorodi, buhă, cea, chefni, chirăi, clefăi, clipoci, colcăi, crîşca, fîrnîi, fojgăi, ghiorţ, hait, harhăt, hălălăi, hîc, hîş, huhura, hurdui, huţa, huzdup, jap, mehăi, mihoho, ocăi, onănăi, pac, pirui, pitpalac, pîc, pîrţ, plici, şîşîi, şovîlc, ştiobîlc, ştiuldic, ticăi, tipa, trosc, ţîţîi, ui, zbîrîi, zîzîi, zumbăi, zumzăi, zvîr | 1 derivat |

Procedînd la fel ca în cazurile anterioare, obţinem tabelul 10.10 şi reprezentarea grafică din figura 10.10, care este şi de această dată tot o dreaptă.

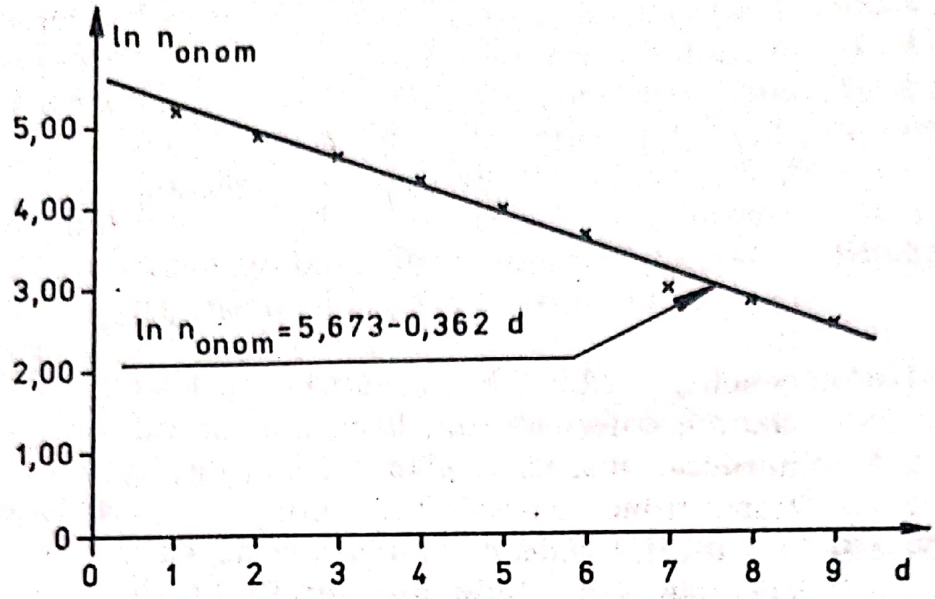
*Tabelul 10.10.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 191 | 140 | 104 | 73 | 51 | 36 | 19 | 16 | 12 |
| In (Nr. cuv.-bază) | 5,25 | 4,94 | 4,64 | 4,30 | 3,93 | 3,58 | 2,94 | 2,77 | 2,48 |



Scanned with OKEN Scanner

FIG. 10.10




În ecuaţia ei:

$$ln\, n_{onom} = 5{,}673 - 0{,}362\, d \qquad (10.10)$$

se remarcă pe de o parte valoarea relativ mică a pantei (inferioară, de exemplu, celei corespunzătoare elementului maghiar = 0,379) şi, pe de altă parte, ecartul mai redus al punctului de abscisă nulă faţă de cel calculat (5,84 faţă de 5,67), ambele indicii ale unei vechimi apreciabile. Dintre toate influenţele studiate, numai fondul de bază daco-roman şi superstratul[1] vechi slav prezintă niveluri mai scăzute ale acestor doi parametri, ceea ce spune multe despre vîrsta medie a cuvintelor onomatopeice.

Egalînd cu zero membrul al doilea al ecuaţiei (10.10), aflăm abscisa punctului de intersecţie a dreptei cu axa orizontală. Aceasta este de cca 15,7. E deci de presupus că cel mai productiv cuvînt onomatopeic va trebui să aibă circa 15—16 derivate. Nu este însă aşa. Campioana onomatopeelor, interjecţia *pic* are nici mai mult nici mai puţin decît 23 de derivate, în contrast flagrant cu cei trei termeni care o secondează în clasament *(balang, fîş, zur)* şi care nu înregistrează decît cîte 11 derivate fiecare. Poziţia cu totul aparte a acestui cuvînt pledează în favoarea marii sale vechimi. Ştim de altfel (v. tabelul 9.1) că, exceptîndu-l pe *pic,* toate cuvintele româneşti ce au dat 20 sau mai multe derivate aparţin fondului primordial al limbii, 17 fiind latine şi

[1] Al. Rosetti, I.L.R. III ed. V, 1964, p. 37.

2 de substrat *(moş, fărîmă)*. Ipoteza că şi *pic* ar putea fi de dată latină (o asemenea onomatopee putea exista deopotrivă în limba băştinaşilor ca şi în latină, unde îl avem pe atît de sugestivul *pîcus* = ciocănitoare) nu trebuie exclusă fără argumente temeinice.

### 10.11. Elementul german

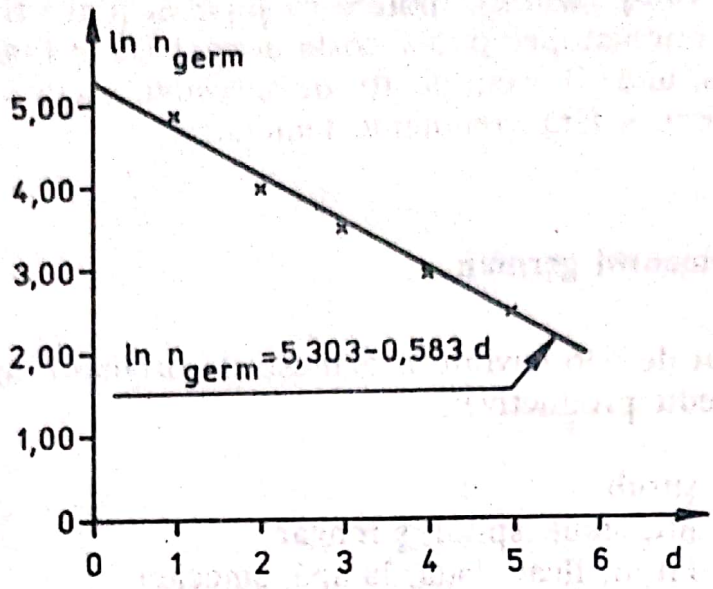
Din totalul de 636 cuvinte româneşti de origine (numai) germană, 124 s-au dovedit productive:

| | | |
|---|---|---|
| 1. | şurub | 11 derivate |
| 2—5. | nit, sfanţ, spital, ştrengar | 6 derivate |
| 6—10. | fălţui, fleac, ladă, lampă, şmecher | 5 derivate |
| 11—17. | lac (=vernis), sorta, şablon, şlefui, şnur, ştanţa, vacs | 4 derivate |
| 18—31. | arest, cocs, iot, pantof, pensulă, struţ (=buchet), şopron, şperţ, şteamp, tîndăli, troacă, turn, valţ, vată | 3 derivate |
| 32—50. | abstract, baiţ, bere, flec, jeţ, malţ, patrulă, referi, rihtui, rochie, şamotă, şerui, şpălţui, ştemui, şvab, trap, ţigară, zeţar, zeţui | 2 derivate |
| 51—124. | abilita, absolvent, absolvi, accidenţe, aclimatiza, baider, blanşir, bonitate, butnar, cartof, chelner, chibiţ, chit, cian, clavir, comasa, component, componist, comporta, corfă, cura (= trata), distona, docent, exmatricula, faliment, felcer, firmă, flaşnetă, gater, gheşeft, glajă, gnoseologie, groş, grund, haltă, hăcui, lector, lombarda, maistru, matriţă, păsui, pergament, pichet (de graniţă, de grevă), politiza, porţelan, preparandie, procent, profesionist, ramă, raşpel, recuzită, repetent, rips, riţui, rutean, scriptic, spiţuri, stroh, şapirograf, şlep, şorţ, şpraiţui, ştemplui, ştrecuri, tabac (= tutun), teoretic, tron (= ladă, sicriu), tuş, ţambal, ţel, ţigaretă, ulan, vană, vest | 1 derivat |



Scanned with OKEN Scanner

FIG. 10.11




Tabelul 10.11. şi fig. 10.11. ţin seama în plus de cele 8 cuvinte cu etimologie dublă ori triplă, dintre care *şină* (germ. & pol.) are 6 derivate, *şanţ* (germ. & pol.) — 5, *cană* (germ. & bg.) — 2, iar *iuncher, maior, ştraf* (toate trei germ. & rus), *ştafetă* (germ. & it.) şi *podzol* (germ. & fr. & rus) — cîte 1 derivat.

*Tabelul 10.11.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nr. cuv.-bază | 132 | 53 | 33 | 19 | 12 |
| ln (Nr. cuv.-bază) | 4,88 | 3,97 | 3,50 | 2,94 | 2,48 |

Există şi aici un decalaj considerabil între primul clasat, substantivul *şurub* şi secondanţii săi, dar, spre deosebire de situaţia anterior analizată a onomatopeei *pic*, productivitatea reală nu o depăşeşte cu mult pe cea scontată, ceea ce arată că nu este vorba de o anomalie. Din ecuaţia dreptei caracteristice, care este:

$$ln\ n_{germ} = 5{,}303 - 0{,}583\ d \qquad (10.11)$$

ar rezulta, pentru cel mai productiv împrumut german, un număr de 9 derivate sau, mai exact, dacă ţinem seama şi de marja de rotunjire



Scanned with OKEN Scanner

a valorilor întregi, de 8,36—10,23 derivate, interval în raport cu care productivitatea reală ($d = 11$) nu prezintă decît un ecart nesemnificativ.

În domeniul termenilor neproductivi, se înregistrează un nou record: faţă de 89, cîţi trebuie să fie conform ecuaţiei (10.11), există nu mai puţin de 530 de cuvinte de origine germană care nu au dat naştere nici unui derivat!

### 10.12. Influenţa ucraineană

Dificultăţi asemănătoare, mutatis mutandis, celor pe care le întîmpină încercările de a discerne net între aporturile lexicale sud-slave (paleo-slav, bulgar, sîrbocroat sau mixt) survin şi în domeniul împrumuturilor din limbile slave răsăritene (ucraineană, polonă, rusă). În general, aportul ucrainian poartă însemnele unei vechimi superioare. Primele contacte lingvistice cu ucrainienii datează din secolele XII—XIII[1], epocă de expansiune românească în nord-estul Europei[2] care a culminat, se pare, cu emigrările din secolele XIV—XV, cînd dreptul „valah" ajunsese să se aplice nu doar în Galiţia (în cca 350 de sate), ci şi în cnezatele Halici, Volhynia şi Kiev[3]. Comparaţia dintre etimologiile propuse de DLRM şi cele avansate în DEX evidenţiază nesiguranţe de diagnostic, depăşite de regulă prin acceptarea provenienţei multiple acolo unde mai demult se indica o origine unică. Iată cîteva exemple:

| DLRM | | DEX |
|---|---|---|
| *plită* | < ucr. *plyta* | rus. ceh. *plita*, ucr. *plyta* |
| *poleac* | < ucr. *poljak* | pol. *polak*, rus. scr. ucr. *poljak* |
| *posteucă* | < ucr. *pidstavka* | bg. rus. *podstavka*, ucr. *pidstavka* |
| *şafăr* | < ucr. *šafer* | germ. *Schaffer*, ucr. *šafer*, pol. *szafar* |

[1] Nu însă înainte de veacul al XII-lea cînd *g*-ul rus comun trece în *h* ucrainian, prezent în cuvintele împrumutate de limba română (substantive comune ca *horn, hulub*, dar şi toponimice: *Horodnic, Hulubăţ*).

[2] În anul 1164 păstori „vlahi" îl capturau la graniţa Galiţei pe viitorul împărat bizantin Andronic Comnen (A. Rosetti I.L.R. III, ed. V, 1964, p. 32).

[3] A. Rosetti, I.L.R. V. ed. V, 1964, p. 117—118.



Scanned with OKEN Scanner

O pulverizare similară a surselor etimologice se întîlneşte şi în domeniul împrumuturilor considerate în DLRM drept polone:

| *DLRM* | | *DEX* |
|---|---|---|
| *castron* | < pol. *castrol* | cf. sas. *Kastrol,* scr. *kastrola,* pol. *kastrol* |
| *palancă* | < pol. *palanca* | tc. pol. *palanka*, magh. *palánk* |
| *erhă* | < pol. *jerha* | magh. *irha,* pol. *jercha* |
| *rabin* | < pol. *rabin* | germ. *Rabbiner,* pol. *rabin* |

sau ruseşti

| *DLRM* | | *DEX* |
|---|---|---|
| *holeră* | < rus *holera* | lat. *cholera,* ngr. *choléra,* ucr. *holera* |
| *ofiţer* | < rus *oficer* | pol. *oficer*, rus *ofiţer,* fr. *officier* |
| *podzol* | < rus *podzol* | fr. *podzole,* germ. *Podsol,* rus *podzol* |
| *silitră* | < rus *selitra* | pol. *salitra*, rus *selitra,* scr. *šalitra* |

Ţinînd seama de precizările DEX, materialul lexical din DLRM cuprinde următoarele 66 de cuvinte productive de origine cert şi unic ucraineană:

| | | |
|---|---|---|
| 1. | calic | 7 derivate |
| 2—6. | bahnă, bortă, dubi, ghili, huţul | 5 derivate |
| 7—13. | buhai, ciubotă, domoli, leah, nămol, tencui, toropi | 4 derivate |
| 14—18. | cruşi, holtei, oboroc, spelcă, tarniţă | 3 derivate |
| 19—36. | barabulă, benchet, boroană, burcă, cîrciog, clăpăug, cotigă, cotonog, crohmală, dohot, droşcă, horn, liur, prihană, pruji, scrabă, zăhăi | 2 derivate |
| 37—66. | balercă, balie, bandură, bezmetic, borti, brăhni, butelcă, covali, drîmbă, gotcă, hlei, holcă, hreapcă, hrincă, hrişcă, lozie, motoaşcă, năjit, odolean, opacină, orcic, pacişe, renie, stoler, şalvir, şleau, tambur, tarac, toval, tumăni | 1 derivat |



Scanned with OKEN Scanner

Un număr de 14 termeni au etimologie dublă ucraineană şi rusă:

| | | |
|---|---|---|
| 1. | tovarăş | 7 derivate |
| 2. | borş | 4 derivate |
| 3—6. | burlac, draniţă, polcovnic, zăton | 2 derivate |
| 7—14. | cneaz, covor, drob, paroi, pihotă, poloboc, stănog, voloc | 1 derivat |

Alţii sînt de provenienţă ucraineană şi polonă (*ghint* cu 3 derivate, *cufăr* şi *ţidulă* cu cîte unul), ucraineană şi bulgară (*scorni* cu 6 derivate, *mălină* cu 2), ucraineană şi sîrbocroată (*laz* şi *şovar,* fiecare cu cîte un derivat), ucraineană şi maghiară (*taşcă* cu 3 derivate). Nu lipsesc nici cuvintele productive cu etimologie triplă: ucr. & pol. & scr. *(poleac)*, ucr. & pol. & rus. *(muscal)*.

Totalizînd întregul fond lexical productiv pentru care s-au propus etimoane ucrainiene, se obţin bilanţul din tabelul 10.12 şi reprezentarea grafică din fig. 10.12.

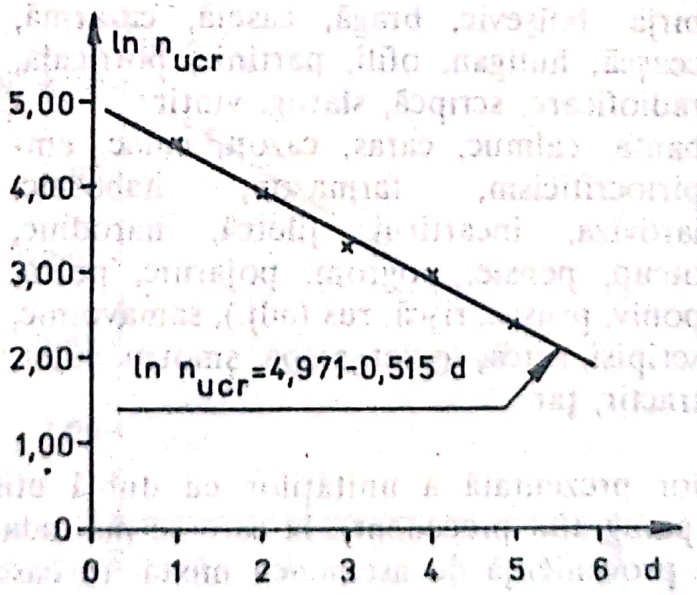
*Tabelul 10.12.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nr. cuv.-bază | 92 | 50 | 27 | 20 | 11 |
| ln (Nr. cuv.-bază) | 4,52 | 3,91 | 3,30 | 3,00 | 2,40 |

Ecuaţia dreptei caracteristice este:

$$ln\, n_{ucr} = 4{,}971 - 0{,}515\, d \qquad (10.12)$$

**FIG. 10.12**






Scanned with OKEN Scanner

Considerînd panta dreptei ca indicator al vechimii, sîntem îndreptăţiţi să atribuim împrumuturilor ucrainiene o vîrstă medie mai mare decît cea a cuvintelor turceşti şi germane, dar inferioară vîrstei aporturilor bulgar sau maghiar. Către o concluzie asemănătoare conduce şi observaţia că în locul a $e^{4,971} - e^{4,356} = 45$ de cuvinte neproductive de origine ucraineană cîte prognozează ecuaţia (10.12), întîlnim în DLRM de peste 4 ori mai multe şi anume 197, ceea ce constituie un semn că peste fondul medieval s-a suprapus şi un strat considerabil de împrumuturi mai recente (probabil sec. XVIII—XIX), majoritatea acestora din urmă rămînînd însă de uz exclusiv regional. De altfel regionalismele reprezintă ceva mai mult de jumătate din totalul cuvintelor româneşti de provenienţă ucraineană din DLRM.

### 10.13. Influenţa rusească

Tabloul productivităţii componentei de origine rusă a lexicului se obţine prin contopirea listei cuvintelor productive cu etimon rusesc unic:

| | | |
|---|---|---|
| 1. | franţuz | 8 derivate |
| 2. | cazac | 5 derivate |
| 3—6. | cavaler, lipovean, mojic, poştă | 4 derivate |
| 7—13. | brac, brişcă, poliţie, raion, rus (subst.), spirt, surtuc | 3 derivate |
| 14—27. | birjă, bolşevic, bragă, casetă, cazarmă, ceaşcă, huligan, ofili, partinic, pătricală, radioficare, scripcă, sfarog, vintir | 2 derivate |
| 28—57. | bantă, calmuc, caras, cazon, culac, empiriocriticism, farmazon, habotnic, iaroviza, încartirui, jiletcă, narodnic, otcup, pensie, pogrom, pojarnic, polcă, poniv, prusac, rişcă, rus (adj.), samavolnic, scripisi, setcă, soviet, stepă, şmotru, şuler, tractir, ţar | 1 derivat |

cu lista anterior prezentată a unităţilor cu dublă etimologie ruso-ucraineană (v. paragraful precedent), la care se mai adaugă un număr de termeni de provenienţă de asemenea mixtă: rusească şi poloneză (*arendă* cu 11 derivate, *şofran* cu 4, *bunt* cu unul singur), rusească şi



Scanned with OKEN Scanner

germană (*iuncher, maior* şi *ştraf* cu cîte un derivat), rusească şi maghiară (*mondir* cu 2 derivate), rusească şi bulgărească (*ispravnic* cu 5 derivate), rusească şi latină savantă (*miliţie* cu 1 derivat), rusească şi franceză (*secţie* şi *ura* cu cîte 1 derivat), rusească, maghiară şi poloneză (*puşcă* cu 13 derivate), rusească, bulgărească, sîrbocroată şi poloneză (*şcoală* cu 13 derivate), rusească, poloneză şi franceză (*ofiţer* cu 6 derivate), rusească şi poloneză (*muscal* cu 5 derivate), rusească, sîrbocroată şi poloneză (*silitră* cu 1 derivat), rusească, franceză şi germană (*podzol* cu 1 derivat).

Situaţia generală a elementului rusesc este reflectată de tabelul 10.13 şi de figura 10.13, iar ecuaţia dreptei caracteristice are forma:

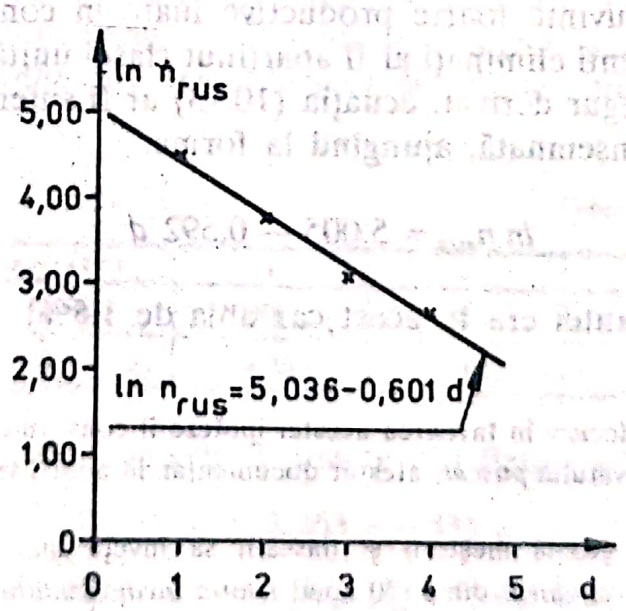
$$\ln n_{rus} = 5{,}036 - 0{,}601\ d \qquad (10.13)$$

*Tabelul 10.13.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nr. cuv.-bază | 90 | 44 | 23 | 15 |
| ln (Nr. cuv.-bază) | 4,50 | 3,78 | 3,14 | 2,71 |

Numărul probabil de derivate ale celui mai productiv împrumut rusesc ar trebui să se situeze, potrivit ecuaţiei 10.13, în jurul cifrei de 8,4. Totuşi trei cuvinte, toate cu etimologie dublă sau triplă, depăşesc acest nivel de productivitate, ceea ce pune sub semnul întrebării

FIG. 10.13






Scanned with OKEN Scanner

provenienţa lor (şi) din limba rusă. Este vorba de *puşcă* şi *şcoală*, termeni cu siguranţă anteriori primelor raporturi lingvistice cu ruşii (cuvîntul *puşcă* desemna în trecut tunul şi e foarte probabil să ne fi venit din maghiară sau/şi polonă încă din secolul al XV-lea, cînd românii iau pentru prima dată contact cu arma respectivă[1]; *şcoală* nu este nici el mult mai recent, întîia sa consemnare în scris[2] datînd din 1570) şi de *arendă*, a cărui explicaţie prin pol. *arenda* este pe deplin satisfăcătoare.

În ce măsură excluderea acestor elemente suspecte influenţează parametrii dreptei caracteristice? Fără ele, tabelul 10.13 capătă forma 10.13.1,

*Tabelul 10.13.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nr. cuv.-bază | 87 | 41 | 20 | 12 |
| ln (Nr. cuv.-bază) | 4,47 | 3,71 | 3,00 | 2,48 |

iar ecuaţia 10.13 devine:

$$ln\ n_{rus} = 5{,}085 - 0{,}668\ d \qquad (10.13.1)$$

Este interesant de observat că panta dreptei a crescut cu cca 11%, deşi unităţile eliminate nu reprezintă decît 3,33% din total. Se evidenţiază astfel dependenţa strînsă a aspectului ecuaţiei caracteristice de numărul de cuvinte foarte productive luate în considerare. Dacă, dimpotrivă, termenii eliminaţi ar fi aparţinut clasei unităţilor ce posedă numai cîte un singur derivat, ecuaţia (10.13) ar fi suferit o modificare mult mai puţin însemnată, ajungînd la forma:

$$ln\ n_{rus} = 5{,}005 - 0{,}592\ d \qquad (10.13.2)$$

Schimbarea pantei era în acest caz abia de 1,5%!

---

[1] Un argument decisiv în favoarea acestei ipoteze îl constituie existenţa încă din sec. al XV-lea a derivatului *puscar*, atestat documentar în anul 1487 (v. Gh. Mihăilă, DLRV, p. 145).

[2] „Aşişderea în şcoală meşterii şi dascălii să înveţe mai vîrtos româneşte" (Epilogul *Octoihului românesc* din 1570 apud *Istoria învăţămîntului din România* vol. 1, Ed. Didactică şi pedagogică, Bucureşti, 1983, pag. 109).



Scanned with OKEN Scanner

## 10.14. Împrumuturi italiene

În ordinea productivităţii lor, cuvintele-bază de origine italiană care au dat derivate în româneşte sînt:

| | | |
|---|---|---|
| 1—2. | amor, distruge | 5 derivate |
| 3—7. | burghez, chinez, italian, student, trata | 4 derivate |
| 8—14. | asedia, balsam, căpitan, dispreţ, reuşi, siloz, trafora | 3 derivate |
| 15—39. | acont, ancoră, bagatelă, barcă, ciment, costa, drit, englez (adj.), fascism, fascist, încasa, medalie, papagal, raşchetă, raţiona, respinge, romanţa, rubin, sconta, sfida, sfrunta, statal, storna, studia, suplini | 2 derivate |
| 40—79. | amic, apolitic, bancă, baston, bilet, calendar, caper, casier, cifră, displăcea, englez (subst.), exemplifica, exploda, gheaţă, gondolă, grosolan, înscena, leal, libret, locotenent, mantelă, monedă, monopolist, mumifica, novelă, restanţă, schiţa, sforţa, siguranţă, solfegia, solo, stampa, stima, sudit, tapet, timpan, tril, valută, verosimil, vinci | 1 derivat |

Adăugînd acestora cuvîntul *spaniol* pe care DLRM îl face să descindă atît din it. *spagnolo* cît şi din fr. *espagnol* şi care are 3 derivate *(spaniolesc, spanioleşte, spanioloaică)* se obţine bilanţul din tabelul 10.14:

*Tabelul 10.14.*

| NR. DERIVATE ≥ | 1 | 2 | 3 |
|---|---|---|---|
| Nr. cuv.-bază | 80 | 40 | 15 |
| ln (Nr. cuv.-bază) | 4,38 | 3,69 | 2,71 |

Dreapta dusă printre cele 3 puncte din figura 10.14 are ecuaţia

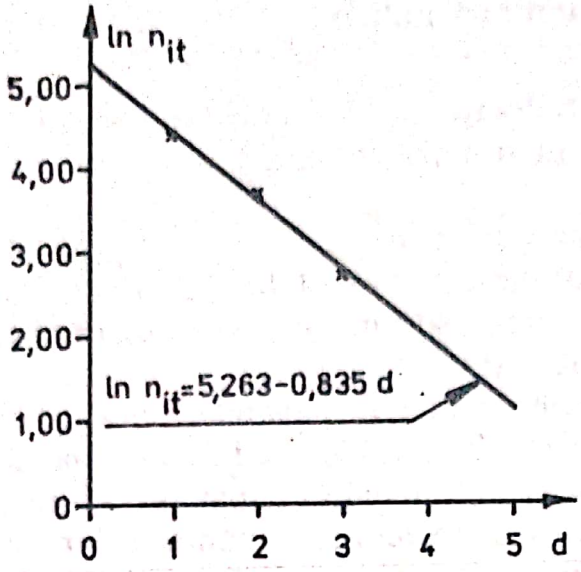
$$\ln n_{it} = 5{,}263 - 0{,}835\ d \qquad (10.14)$$

Panta ei este, după cum se vede, superioară tuturor celor întîlnite pînă acum ($m = 0{,}830$), ceea ce ar putea însemna că, în raport cu



Scanned with OKEN Scanner

Fig. 10.14




influenţele discutate anterior, cea italiană este cea mai recentă. Să nu uităm însă de existenţa şi a unor cuvinte, în general mai vechi (stă mărturie productivitatea lor ridicată), a căror provenienţă italiană este nesigură, dar posibilă: *alamă* cf. it. *lama* cu derivatele *alămar*, *alămărie*, *alămi*, *alămire*, *alămit*, *alămiu; cicală* cf. it. *cicala* cu derivatele *cicăleală*, *cicăli*, *cicălire*, *cicălitor; ciung* cf. it. *cionco* cu derivatele *ciungar*, *ciungări*, *ciungi; ciot* cf. it. *ciotto* cu derivatele *cioată*, *ciotîrcă*, *ciotur*, *cioturar*, *ciuteică; piţiga* cf. it. *pizzicare* cu derivatele *piţigăia*, *piţigăială*, *piţigăiat*, *piţigoi*, *piţiguş*.

Dacă se ţine seama şi de aceşti termeni, tabelul 10.14 ia forma 10.14.1.

*Tabelul 10.14.1.*

| NR. DERIVATE ≥ | 11 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nr. cuv.-bază | 85 | 45 | 20 | 11 |
| ln (Nr. cuv.-bază) | 4,44 | 3,81 | 3,00 | 2,40 |

iar ecuaţia dreptei caracteristice devine:

$$ln\, n_{it} = 5{,}146 - 0{,}693\, d \qquad (10.14.1)$$

Redusă cu aproape 17%, panta dreptei a ajuns astfel practic egală cu cea aferentă aportului turcesc. Centrul de greutate al înrîuririi italiene a fost împins astfel pînă undeva prin a doua jumătate a secolului al XVIII-lea, datare puţin forţată, în ciuda existenţei atestate a unor contacte lingvistice chiar anterioare anului 1700 (v. capitolul 12).



Scanned with OKEN Scanner

## 10.15. Influenţa latină savantă

Lista celor mai productive cuvinte latine pătrunse în româneşte pe cale cultă se deschide cu controversatul *popor*. Cu această formă şi cu înţelesul de „mulţime a enoriaşilor unei parohii" el este întîlnit încă din 1525 într-un text slavon[1], iar în româneşte apare pentru prima oară[2] în *Pravila sfinţilor apostoli* din 1570. În ciuda dificultăţilor fonetice (mutarea accentului şi transformarea lui *u* în *o*), lingvistica mai nouă (I. Coteanu şi M. Sala[3], Gh. Ivănescu[4]) îl consideră moştenit din lat. *populus,* dar modelat semantic sub efectul unor influenţe mai recente. Grefarea de sensuri neologice pe cuvinte vechi moştenite este de altfel un fenomen destul de frecvent. Să amintim numai exemplul lui *tramă*, continuator al lat. *trama,* dar cu sensuri adăugate după fr. *trame*. În acest din urmă caz, aspectul cuvîntului era de aşa natură încît acţiunea legilor fonetice nu l-a afectat suficient pentru a ne permite să distingem dacă este vorba de un termen moştenit sau de un neologism (apariţia *ă*-ului final este nepertinentă căci ea are loc şi în împrumuturile recente). O invarianţă fonetică asemănătoare (de dată aceasta chiar perfectă) face dificilă identificarea provenienţei în româneşte a substantivului *aer,* ce poate fi deopotrivă urmaşul direct al omonimului său latin ca şi un împrumut mai nou (rudele sale *aerisi* şi *aera* provin unul din neogreacă, altul din franceză).

Situaţii de echivoc etimologic de felul ultimei citate sînt însă rare. În schimb, identificarea corectă a sursei multor neologisme rămîne o problemă dificilă. DEX dă dovadă în această privinţă de mai multă prudenţă decît DLRM, trimiţînd frecvent la mai multe limbi acolo unde anterior se indicase o obîrşie unică. Recunoaştem aici un scrupul ştiinţific lăudabil, dar poate şi o teamă exagerată de asumare a răspunderii diagnosticului, care, din perspectiva cercetării întreprinse de noi, prezintă serioase inconveniente. Practica dublării aproape automate prin termenul francez corespunzător a etimonului latin literar (şi vice-versa) pentru majoritatea neologismelor anulează practic distincţia, ce trebuie oricum făcută, dintre latinisme şi împrumuturile din limbi moderne, confuzie cu urmări nu numai asupra istoriei limbii, ci şi a culturii româneşti în general. Faptul de a şti

---

[1] Gh. Bolocan (red. resp.) *Dicţionarul elementelor românesti din documentele slavo-române,* Bucureşti, 1981.

[2] O. Densusianu *Istoria limbii române* vol. II, pag. 297.

[3] *Etimologia si limba română* Ed. Acad. R.S.R., pag. 90—93.

[4] *Istoria limbii române* Ed. Junimea 1980, pag. 245.



Scanned with OKEN Scanner

exact ce cuvinte sînt preluări cărturăreşti din lexicul latin literar prezintă o importanţă ce depăşeşte sfera lingvistică propriu-zisă, întrucît el contribuie la clarificarea raporturilor dintre cultura noastră şi umanismul european de expresie latină din post-Renaştere. De aceea am socotit necesar să menţinem pe cît posibil distincţia dintre elementul latin savant şi cel romanic acceptînd (în bloc) etimologiile propuse de DLRM (chiar atunci cînd nu eram de acord cu ele[1]). Precizarea este necesară deoarece într-un alt caz, cel al limbilor slave, procedasem altminteri, preferînd originile indicate în DEX.

Iată acum lista, apud DLRM, a fondului de cuvinte româneşti productive de provenienţă latină savantă:

| | | |
|---|---|---|
| 1. | popor | 9 derivate |
| 2—3. | doctor, linie | 8 derivate |
| 4—6. | compune, popula, punct | 7 derivate |
| 7. | rege | 6 derivate |
| 8—11. | coroană, daună, murmura, ştiinţă | 5 derivate |
| 12—19. | leu, premiu, prepune, primar, sigila, stabili, sumă, vistier | 4 derivate |
| 20—38. | audia, convinge, corupe, defect, defrauda, fruct, latin, nutri, odor, pauper, produce, propune, protege, purpură, regula, regulă, satisface, spaţiu, susţine | 3 derivate |
| 39—87. | abrevia, adnota, cauză, circumscrie, clarifica, concepe, conduce, contrage, culpă, deduce, depune, despera, devota, | |

[1] De ce am accepta, de pildă, că rom. *efect* provine din fr. *effet* mai degrabă decît din lat. lit. *effectus* sau rom. *centru* din fr. *centre* (pronunţat *sătr*), iar nu direct din lat. lit. *centrum?* Exemplele se pot lesne înmulţi. Dacă rom. *actor*, în ciuda aspectului grafic identic cu al termenului originar lat. lit. *actor,* prezintă accentul francezului *acteur* din care probabil derivă (în treacăt fie zis, Eminescu pronunţa *áctor*), nu există în schimb nici un motiv de a atribui origine franceză unor cuvinte ca *fáctor, diréctor,* (francez este în schimb adjectivul *directór*), *inspéctor* care pot cel mult proveni din germană. E greu de înţeles opţiunea DLRM pentru etimologii divergente în cazuri ca: *doctor* < lat. lit. *doctor*, *lector* < germ. *Lektor, rector* < fr. *recteur* cînd toţi aceşti termeni prezintă exact aceleaşi caracteristici (pronunţie identică în română, latină şi germană, rostire şi accent diferite în franceză). De asemenea, nu pricepem de ce *compune* şi *produce* sînt explicate prin lat. lit. *componere, producere* în timp ce pentru *dispune* şi *introduce* se dau ca surse fr. *disposer, introduire,* cînd latina poseda atît pe *disponere* cît şi pe *introducere.*



Scanned with OKEN Scanner

dizolva, elibera, emite, exclude, extrage, fortifica, gumă, include, inflama, invada, istorie, matur, nerv, opune, oră, ordin, ordine, pagină, perfect, pondere, predica, prefect, pretor, principiu, proba, promite, promova, pronunţa, redacta, reduce, renega, rezolva, scripte, traduce, trăda, voma — 2 derivate

88—183. aproba, cancelar, capsă, castra, chirurg, cisternă, comite (vb.), compactor, comuta, conspecta, corigent, cretă, cronică, decan, decembrie, deifica, descinde, dirigui, diurnă, divide, excomunica, exercita, expune, extrage, frecventa, germen, germina, gerunziu, gigant, himeră, induce, insulă, interpune, invidie, irupe, încarcera, latifundiu, martiriu, matricolă, medicină, migra, multiplica, muzică, nasicorn, natural, nega, oferi, omite, onest, particular, percepe 1, percepe 2, peregrin, pericol, pertracta, pictor, plus, prefectură, presta, prezbiter, profund, provincie, publica, raţiune, război, recepta, reduplica, remite, repauza, reproba, ricin, rută (bot.), sacrifica, saluta, salva, selbă, simplifica, sista, solvi, specie, spera, spirit, stabil, tabelă, templu, tezaur, titlu, triumfa, triumfal, uman, unic, uniune, unisilab, uzita, vocală, vomita — 1 derivat

Eliminîndu-l, din motivele arătate, pe discutabilul *popor,* şi adăugînd 2 cuvinte cu etimologie dublă: *staţie* (lat. sav. & fr) şi *miliţie* (lat. sav. & rus) ambele cu cîte un derivat (*substaţie, miliţian*) se obţine bilanţul din tabelul 10.15.

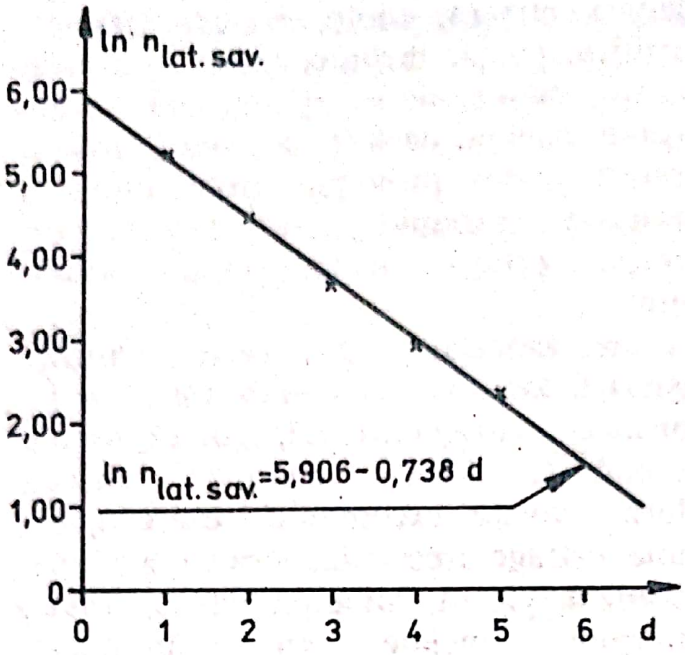
*Tabelul 10.15.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nr. cuv.-bază | 184 | 86 | 37 | 18 | 10 |
| ln (Nr. cuv.-bază) | 5,21 | 4,45 | 3,61 | 2,89 | 2,30 |



Scanned with OKEN Scanner

FIG.10.15




Dreapta caracteristică rezultată (fig. 10.15) are ecuaţia:

$$n_{lat.sav} = 5{,}906 - 0{,}738\ d \qquad (10.15)$$

În afara pantei, superioară ca valoare celei corespunzătoare influenţei turceşti, ruseşti sau germane, discrepanţa dintre numărul de cuvinte neproductive prevăzut de ecuaţie ($e^{5,164} = 175$) şi cel real (= 623) dezvăluie caracterul predominant modern al înrîuririi latine savante asupra vocabularului românesc.

## 10.16. Influenţa „franceză"

Continuăm să menţinem ghilimelele ca semn avertizator asupra felului discutabil în care au fost grupate laolaltă cuvintele socotite în DLRM drept avînd origine franceză[1].

---

[1] S-a observat (v. Sanda Reinheimer Rîpeanu: *Cuvinte de origine franceză?* SCL XXXVIII nr. 2 martie — aprilie 1987, pp. 110—114) că unora dintre ele li s-au atribuit etimoane de-a dreptul inexistente (*motorette*, sursa propusă a rom. *motoretă* nu figurează în nici un dicţionar francez), iar altele posedă sensuri total deosebite de cel din pretinsa limbă-sursă (în franceză *cryptique* înseamnă „care zace în criptă", înţeles complet diferit de al românescului *criptic*).



Scanned with OKEN Scanner

Printre cele 15 121 de cuvinte primare (nederivate) de origine „franceză“ din DLRM, 2101 sînt productive. Enumerarea lor fiind fastidioasă şi, mai ales, devoratoare de spaţiu tipografic, ne limităm la consemnarea listei celor care au generat în româneşte cîte cel puţin 3 derivate:

| | | |
|---|---|---|
| 1. | semna | 18 derivate |
| 2. | alimenta | 8 derivate |
| 3—6. | aprecia, lustru, valoare | 7 derivate |
| 7—10. | califica, pachet, soldat, şosea | 6 derivate |
| 11—25. | angaja, anima, dispune, evalua, fabrica, guverna, imprima, modela, mutila, pava, planta, proprietar, satura, slavon, zinc | 5 derivate |
| 26—63. | aglomera, bobina, calcina, completa, comprima, concentra, condiţiona, congestiona, corect, decepţie, determina, fasona, fracţiona, gaz, gaza, gazetă, grad, incarna, încadra, încînta, învesti, lamina, marca, militar, mina, nobil, normă, portret, presa, program, proiect, ramoli, rarefia, regal, selecţiona, suda, umili, vers | 4 derivate |
| 64—203. | absorbi, achiziţie, actor, actualiza, adapta, administra, afirma, ajusta, amoreză, anchetă, articula, artificial, asfalta, ataca, ataşa, balansa, bandaja, broşa, calcula, calibru, canar, categorie, compromite, conserva, considera, contur, copia, crom, cultural, dactilografia, decupa, defini, definitiv, demasca, demobiliza, denunţa, deranja, deriva, descuraja, detaşa, devasta, dezinteresa, dig, dilata, dirija, disocia, doctrină, doliu, domestic, duel, echilibra, economic, estima, evidenţă, examina, exploata, fila, finanţa, flagela, forma, fotografia, funcţionar, funda, furie, gimnaziu, grima, hipnotiza, ilegal, ilumina, ilustra, impune, incendia, incrimina, inginer, intitula, introduce, invita, izola, înclina, încorpora, încuraja, înhuma, licita, limita, mască, mecanic, mecaniza, miner, mobilă, mobiliza, moderniza, naţional, nivela, nota, numerota, obiectiv, | |



Scanned with OKEN Scanner

ofensa, pagina, parveni, patina, perforaj, placa, plana, planifica, politic, poliza, populariza, precipita, preşedinte, protesta, rafina, raţionaliza, reclama, recrea, refugia, registru, regla, sciziune, sculpta, secţiona, sirop, stiliza, sublinia, subordona, substitui, surprinde, şantaj, şicana, tehnic, tigru, tip, tortura, trasa, travesti, trefila, tricota, tub, unifica, utila, vulcaniza 3 derivate

Examinarea listei reliefează caracterul eteroclit al materialului lexical „francez". Cu precădere printre rangurile superioare figurează destui termeni a căror provenienţă franceză este discutabilă. Chiar în capul listei întîlnim un cuvînt de o productivitate cu totul suspectă, aflată în vădit contrast cu cea a restului unităţilor „franceze", verbul *semna,* din care descind nu mai puţin de 18 cuvinte: *însemna, însemnare, însemnat, însemnărică, însemnătate, însemnător, consemna, consemnare, contrasemna, contrasemnare, contrasemnătură, neînsemnat, semnare, semnat, semnătură, subsemna, subsemnat, subsemnătură.* Dicţionarul indică drept etimon verbul francez *signer*, urmat de o paranteză în care se precizează „după *semn*". E vorba, precum se vede, nu de un împrumut, ci de un calc lexical. Mai corect, după opinia noastră, ar fi fost ca în loc de „din fr. *signer* (după *semn*)" să se menţioneze, exact dimpotrivă, „din *semn* (după fr. *signer*)", aşa cum s-a procedat, de altfel, în cazul unor calcuri ca *sinucide* („din *sine* + *ucide* (după fr. *suicider*)" sau *înlocui* („din *în* + *loc* (după fr. *remplacer*)"). La fel, era normal ca rom. *încînta* să fie explicat ca alcătuit din *în* + *cînta* după modelul fr. *enchanter*, iar nu ca preluat direct din franceză, cum lasă să se înţeleagă dicţionarul. Socotim chiar că şi semicalcurile ar trebui tratate într-un mod asemănător, pentru a le distinge de împrumuturile propriu-zise (DLRM indică exact acelaşi etimon, şi anume fr. *consanguin,* atît pentru rom. *consîngean* cît şi pentru rom. *consanguin*, deşi deosebirea formală netă dintre cele două cuvinte româneşti ar fi reclamat o minimă explicaţie).

Dilema francez/latin savant este reprezentată mai ales prin verbe: rom. *anima* < fr. *animer*/lat. lit. *animare,* rom. *mutila* < fr. *mutiler*/lat. lit. *mutilare,* rom. *planta* < fr. *planter*/lat. lit. *plantare,* rom. *satura* < fr. *saturer*/lat. lit. *saturare,* rom. *dispune* < fr. *disposer*/lat. lit. *disponere,* toate spicuite dintre cele mai productive 25 de cuvinte. Nu lipsesc însă nici substantivele susceptibile de o dublă explicaţie etimologică: rom. *lustru* < fr. *lustre*/lat. lit. *lustrum* cu 7 derivate, rom. *proprietar* < fr. *proprietaire*/lat. lit. *proprietarius* cu 5 derivate, rom. *grad* < fr. *grade*/lat.



Scanned with OKEN Scanner

lit. *gradus* şi rom. *vers* < fr. *vers*/lat. lit. *versus,* fiecare cu cîte 4 derivate. Am enumerat într-o notă de sub 10.15 şi o parte (redusă) dintre cazurile în care credem că alternativa ar trebui tranşată în favoarea provenienţei latine literare.

Avînd însă în vedere că ne-am propus încă dintru început să ne limităm la indicaţiile dicţionarului, rezistînd tentaţiei de a avansa soluţii etimologice proprii, am preluat tale-quale materialul lexical considerat de DLRM drept împrumutat din franceză, distanţîndu-ne de poziţia autorilor numai prin plasarea sistematică între ghilimele a indicaţiei privind proveniența cuvintelor respective.

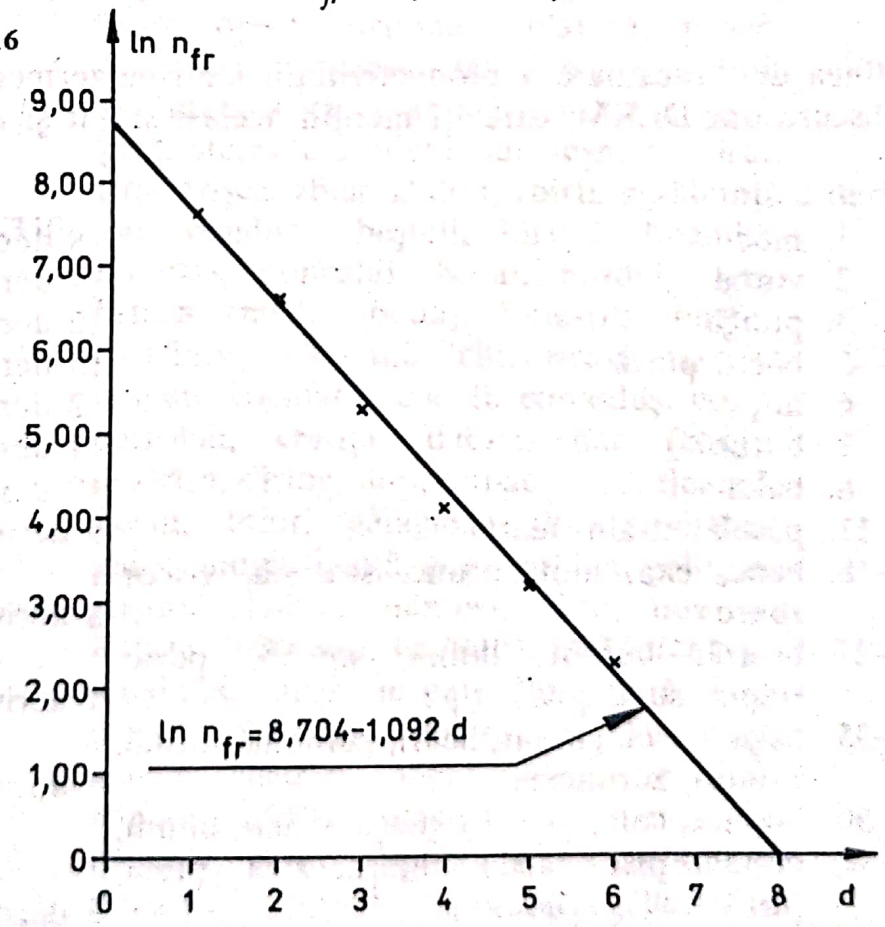
Tabelul 10.16, figura 10.16 şi ecuaţia 10.16 exprimă sintetic bilanţul influenţei „franceze" asupra vocabularului limbii române pînă în jurul anului 1950.

*Tabelul 10.16.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Nr. cuv-bază | 2102 | 763 | 203 | 63 | 25 | 10 |
| ln (Nr. cuv.-bază) | 7,65 | 6,64 | 5,31 | 4,14 | 3,22 | 2,30 |

$$ln\ n_{fr} = 8{,}704 - 1{,}092\ d \qquad (10.16)$$

FIG. 10.16






Scanned with OKEN Scanner

Bănuiala noastră că fruntaşul clasamentului, verbul *semna*, nu are ce căuta printre împrumuturile din franceză este confirmată de observaţia că intersecţia dreptei (10.16) cu axa absciselor are loc în dreptul valorii $d = 7{,}97 \approx 8$, dovadă că cel mai productiv cuvînt provenit din franţuzeşte a produs în jur de 8 derivate (este cazul lui *alimenta*), iar nicidecum 18 (!) cîte posedă cuvîntul incriminat.

În ceea ce priveşte corelaţia dintre vechimea medie şi pantă, trebuie spus că înclinarea dreptei 10.16 este cea mai mare întîlnită de noi şi ea ar putea fi chiar mai mare în ipoteza că o serie de cuvinte productive, dintre care pe unele le-am menţionat mai sus, s-ar transfera elementului latin savant. Ca în cazul tuturor influenţelor mai noi, numărul de cuvinte neproductive (= 13.020) depăşeşte cu mult nivelul prognozat de ecuaţia dreptei caracteristice (≈ 4 000). O bună parte a împrumuturilor din franceză nu sînt deci încă suficient asimilate pentru a-şi fi putut întemeia o familie lexicală proprie.

### 10.17. Cuvinte cu etimologie necunoscută

În ordinea descrescătoare a productivităţii, ierarhia termenilor de origine obscură din DLRM, care îşi menţin acelaşi statut şi în DEX este:

| | | |
|---|---|---|
| 1. | moc | 18 derivate. |
| 2. | vorbă | 17 derivate |
| 3. | pungă | 16 derivate |
| 4—5. | băiat, pînză | 15 derivate |
| 6. | hoţ | 12 derivate |
| 7. | butuc | 11 derivate |
| 8. | baltă | 10 derivate |
| 9—11. | păcală, străin, tearfă | 9 derivate |
| 12—18. | bade, ban, moţ, prunc, scălîmb, viscol, zburda | 8 derivate |
| 19—27. | brînză, burduf, doină, migăli, perie, stăpîn, sută, şovăi, ţipa | 7 derivate |
| 28—35. | băga, burtă, ghiont, licări, paragină, ticală, zbînţui, zdruncina | 6 derivate |
| 36—50. | buimac, colb, ghes, gogoaşă, mălai, mîhni, motan, pisc, răbda, reteza, saia, ţugui, ţurcă, vuvă, zgudui | 5 derivate |



Scanned with OKEN Scanner

51—81. afin, aromi, băţ, bîrfi, bot, boţ, bruftui, căpută, cruţa, desfăta, domol, fulg, hartan, îngăima, mămăligă, nasture, pardosi, pipernici, răsură, refeca, rezema, scăpăra, stingher, stîncă, şubred, tiv, ţîră, uimi, urdă, zăpăci, zdreli — 4 derivate

82—108. baci, boarfă, breabăn, bulgăre, cîrlan, crîncen, destrăbăla, dibui, fercheş, ghici, gogon, închiondora, încotoşmăna, jupîn, maldăr, năbădăi, neghină, prăvălie, rostogol, stîrni, strugure, terci, ţurţur, zer, zgău, zorzoane — 3 derivate

109—167. adulmeca, aoleu, bălăbăni, bîrzoi, bleojdi 1, bleojdi 2, borî, brănui, brudiu, chirci, chiţibuş, cinchi, cintez, cioareci, cîrlionţ, cîrti, clinti, cotropi, dăula, deşănţa, glonţ, gog, jarcă, lăbărţa, lehăi, leu, lihni, mîglisi, moşmoli, nărui, neghiob, oblon, olori, opacă, ortoman, platoşă, posac, praftură, prepeleac, răsti, răşlui, rătuti, săhăidac, scociorî, smidă, sturluiba, şiroi, şonţ, şterpeli, ştiulete, şui, teşmeni, ticăi, tîrşi, trişcă, zbici, zbihui, zbîrli, zgrăbunţă — 2 derivate

168—310. aui, bandură, beşteli, bîrliga, boambă, bolboşa, boncălui, borăi, bordei, borţ, brîncă (med.), buhaş, buiestru, burfui, cărpănos, chelfăni, chirfoseală, cimili, ciuguli, cocioabă, cocoli, corcoduş, corlă, cotrobăi, cratiţă, dabilă, dar (conj.), dezbăra, dîrlog, dup, dupac, durd, fleancă, flecui, fuior, gălăgie, ghiborţ, ghiftui, ghioc, giugiuli, gîţă, gogi, goldan, gologan, goştină, hăbuc, hămuşi, hîrşi, hudubă, hutupi, îmbodoli, închelba, închiorchioşa, janghină, japşă, jerpeli, jip, lastră, laţe, lespede, lolotă, lostropan, marţafoi, măceş, mădări, melc, mite, modîrlă, năgară, năplăi, olastisi, pas (unealtă), părădui, piguli, pistosi, pîlnie, poloboc, pongosi, poznă, presură, preveşti, promoară, pruhă, pupui, puşlama, rapăn,



Scanned with OKEN Scanner

răfui, răpciugă, rediu, retevei, rîncă, rîznă, sclipui, scrînciob, sfîrc, sfîrlă, sfîrloagă, sfriji, sitar, siteav, spînz, sprăhui, sprinţar, stinghie, stînă, stogoş, strai, şoroboti, ştează, tăbîrcă, tăşulă, ticlui, titiri, tîromeală, tîrş, tolăni, tolocăni, toroipan, trandada, troinic, tulnic 1, tulnic 2, ţanţoş, ţîrîă, ţîrţîră, ugui, vijelie, vitrui, vîlced, vîlvoi, vîndălac, vîrcă, zaică, zbrehui, zbughi, zeberi, zgancă, zgîmboi, zgîndări, zgorni, zgrepţăna, zoală, zvecui — 1 derivat

Obîrşia acestor cuvinte este extrem de controversată. Pentru unele dintre ele, simpla enumerare a etimologiilor propuse de-a lungul timpului de diferiţi lingvişti români sau străini ar umple pagini întregi. Singurul lucru ce se poate afirma cu certitudine este că originile lor sînt foarte variate, cîteva fiind neîndoielnic străvechi, după cum altele pot aparţine unor straturi lexicale relativ recente. Criteriul pantei ca indicator al vîrstei medii a materialului lexical ne va putea da o idee despre raportul cantitativ dintre cele două categorii de unităţi.

Reprezentate grafic (fig. 10.17.1), valorile din tabelul 10.17.1. conduc la un şir de puncte a căror coliniaritate lasă mult de dorit, urmare evidentă a caracterului compozit al materialului utilizat.

*Tabelul 10.17.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 310 | 167 | 108 | 81 | 50 | 35 | 27 | 18 | 11 |
| ln (Nr. cuv.-bază) | 5,74 | 5,12 | 4,68 | 4,39 | 3,91 | 3,56 | 3,30 | 2,89 | 2,40 |

Metoda celor mai mici pătrate ne îngăduie să trasăm o dreaptă de ecart minim printre aceste puncte. Ecuaţia ei este:

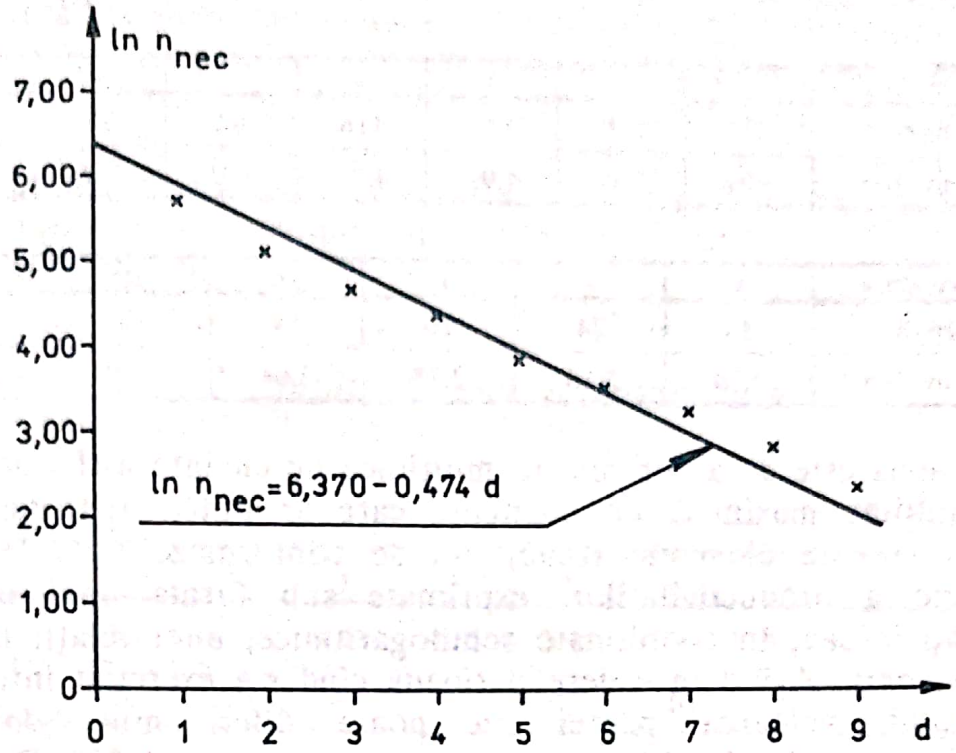
$$ln\, n_{nec} = 6{,}370 - 0{,}474\, d \qquad (10.17.1)$$

Panta dreptei (m = 0,474) se dovedeşte intermediară între cea corespunzătoare aporturilor lexicale vechi (autohton 0,196, latin 0,230, vechi slav 0,312, bulgar 0,362, onomatopeic 0,362, maghiar 0,379) şi a celor moderne (german 0,583, rusesc 0,591, turcesc 0,690, latin savant 0,738, italian 0,830, francez 1,092). Cel mai mult se aproapie ea de valoarea aferentă influenţei greceşti (m = 0,501), cu care se şi



Scanned with OKEN Scanner

Fig. 10.17.1




aseamănă sub aspect diacronic, deoarece aceasta din urmă s-a constituit la rîndul ei din straturi diferite, desfăşurate în timp de-a lungul întregii istorii a limbii române. În orice caz, diferenţa importantă de comportament statistic dintre cuvintele cu etimologie necunoscută şi fondul lexical primordial al limbii este de natură să descurajeze tentativele de a atribui în mod abuziv substratului cea mai mare parte a termenilor a căror sursă nu a fost încă identificată. Este firesţe posibil, ba chiar foarte probabil, ca printre aceşti termeni să se afle într-adevăr şi unele cuvinte provenite din limba traco-dacilor, dar numărul lor nu poate fi prea mare. Dispunem în prezent de suficiente elemente pentru a încerca să stabilim o limită superioară a acestui număr.

Cumulînd aportul cunoscut (la momentul actual) al substratului (v. paragraful 10.2. şi tabelul 10.17.1) cu materialul lexical de origine obscură, ceea ce revine la însumarea tabelelor 10.17.1., extins pentru valorile $9 \leq d \leq 13$, şi 10.17.2., se obţin bilanţul din tabelul 10.17.3. şi reprezentarea grafică din figura 10.17.2.

*Tabelul 10.17.2.*

| $d \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_a$ | 62 | 51 | 39 | 35 | 32 | 24 | 19 | 15 | 13 | 8 | 7 | 7 | 5 |



Scanned with OKEN Scanner

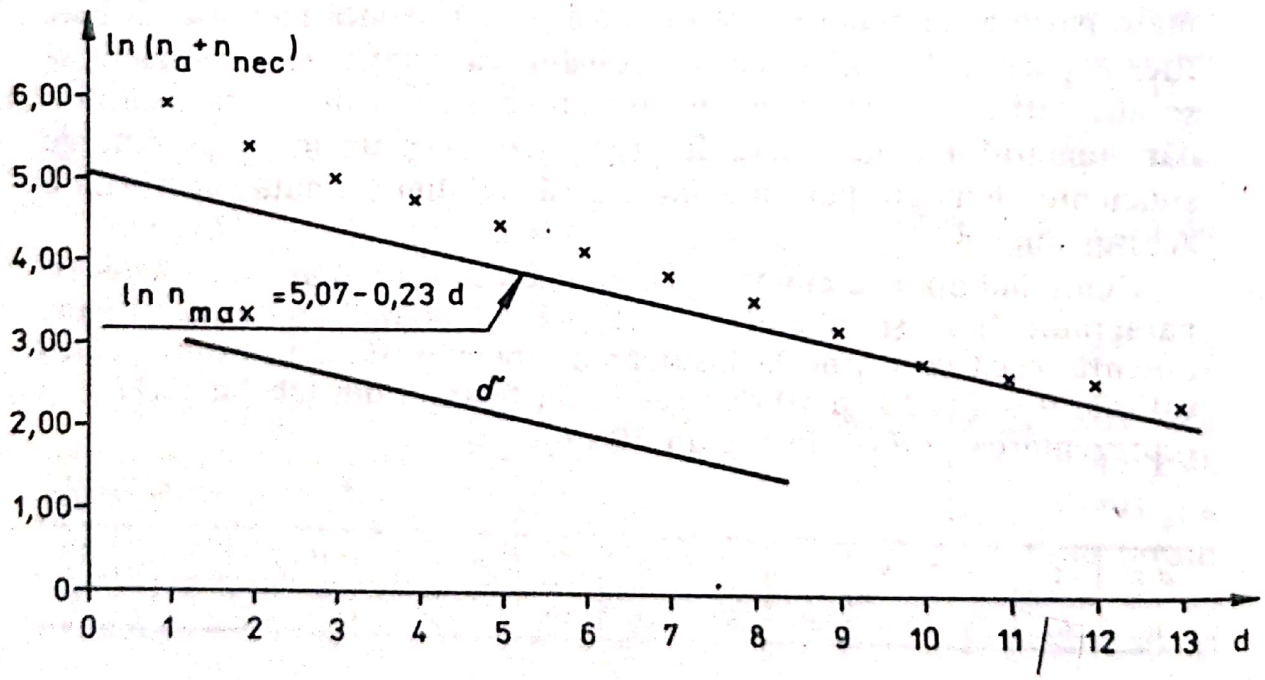
*Tabelul 10.17.3.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 372 | 219 | 147 | 116 | 82 | 59 | 46 |
| ln (Nr. cuv.-bază) | 5,92 | 5,39 | 4,99 | 4,75 | 4,41 | 4,08 | 3,83 |

| NR. DERIVATE ≥ | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 33 | 24 | 16 | 14 | 13 | 10 |
| ln (Nr. cuv.-bază) | 3,50 | 3,18 | 2,77 | 2,64 | 2,56 | 2,30 |

Problema este de a delimita în mulţimea de cuvinte astfel definită o submulţime maximală de elemente care ar putea aparţine substratului. Aceste elemente trebuie să se conformeze unei legi de distribuţie a productivităţilor, exprimate sub forma unei ecuaţii exponenţiale sau, în coordonate semilogaritmice, unei relaţii liniare de pantă dată. Avînd în vedere perioada cînd s-a exercitat influenţa substratului, valoarea pantei nu poate diferi mult de cea corespunzătoare fondului latin moştenit, care este m = 0,230. Dreapta caracteristică ce modelează matematic aportul autohton va trebui să respecte deci această pantă şi să nu depăşească nicăieri, pe întregul său parcurs, cotele punctelor aferente materialului lexical cumulat (de

**FIG. 10.17.2**



origine necunoscută + comun cu albaneza). În termeni geometrici, rezolvarea problemei presupune trasarea pe graficul din figura 10.17.2 a unei drepte de pantă $m = 0{,}230$, situate cît mai sus posibil, fără însă ca vreunul dintre punctele $A_i$ ($i = 1{,}13$) să coboare sub ea. Fiind dată deci o linie $\delta$ cu înclinarea respectivă, dreapta caracteristică a fondului autohton maximal va fi paralela dusă la $\delta$ prin punctul cel mai apropiat de $\delta$, care este cel corespunzător abscisei $d = 10$.

Ecuaţia ei este:

$$\ln n_{max} = 5{,}07 - 0{,}23\, d \qquad (10.17.2)$$

Relaţia (10.17.2) furnizează pentru valorile $d = 1, 2, \ldots, 13$ ale argumentului rezultatele din tabelul 10.17.4.

*Tabelul 10.17.4.*

| $d \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $\ln n_{max}$ | 4,84 | 4,61 | 4,38 | 4,15 | 3,92 | 3,69 | 3,46 | 3,23 | 3,00 |
| $n_{max}$ | 126 | 100 | 84 | 71 | 50 | 40 | 32 | 25 | 20 |

Pentru a uşura comparaţia, să alăturăm liniile finale ale tabelelor 10.17.4. şi 10.17.2, efectuînd totodată scăderea $(n_{max} - n_a)$ (tabelul 10.17.5):

*Tabelul 10.17.5.*

| $d \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $n_{max}$ | 126 | 100 | 84 | 71 | 50 | 40 | 32 | 25 | 20 |
| $n_a$ | 62 | 51 | 39 | 35 | 32 | 24 | 19 | 15 | 13 |
| $n_{max} - n_a$ | 64 | 49 | 45 | 36 | 18 | 16 | 13 | 10 | 7 |

În tabelul 10.17.5. linia $n_{max}$ conţine numărul probabil de cuvinte de substrat, pînă la o productivitate dată, pe care le poate conţine vocabularul limbii române, iar linia $n_a$ — numărul cuvintelor deja identificate drept autohtone prin procedeul comparării cu limba albaneză. Diferenţa lor ne dă răspuns la întrebarea: „cîte dintre cuvintele cu etimologie necunoscută pot aparţine substratului?“

Dacă ne referim la totalul cuvintelor productive (coloana 1), compararea tabelelor 10.17.1. şi 10.17.5. ne arată că, printre cele 310 cuvinte-bază de origine neidentificată, *cel mult* 64, adică circa o



Scanned with OKEN Scanner

cincime, pot proveni din substrat. Ştiut fiind faptul (v. paragraful 10.2) că, spre deosebire de influenţele mai recente, lexicul de substrat respectă, cu o abatere mai mică de 15%, ecuaţia dreptei caracteristice, chiar şi pentru valoarea $d = 0$ a argumentului, ne este uşor să calculăm şi numărul total de cuvinte (productive şi neproductive) de origine obscură ce pot avea obîrşie autohtonă.

Pentru $d = 0$, avem $n_{max} = exp\ 5{,}07 = 159$. Majorînd acest număr cu 15%, obţinem $n_{max} = 183$ cuvinte de substrat, dintre care 92 sînt deja cunoscute (v. paragraful 10.2). Urmează că din totalul celor 1508 cuvinte de provenienţă necunoscută numai *cel mult* 91 sînt autohtone. În ceea ce priveşte unităţile neproductive, raportul este şi mai dezavantajos, căci, avînd în vedere rezultatul anterior, numai $91 - 64 = 27$ de cuvinte din 1198 pot proveni din substrat. Şansele maxime de a fi autohtoni le au termenii cei mai productivi şi în această privinţă putem stabili o ierarhie de priorităţi.

În conformitate cu cifrele din tabelul 10.17.5., printre cele 11 cuvinte cu etimologie necunoscută care au generat în româneşte 9 sau mai multe derivate, un număr de 7 ar putea proveni din substrat. Iată lista celor 11 termeni vizaţi: *moc, vorbă, pungă, băiat, pînză, hoţ, butuc, baltă, păcală, străin, tearfă.* Concluzia noastră, bazată exclusiv pe considerente de ordin matematic, este confirmată de opiniile unor lingvişti, care apreciază drept probabil autohtone cel puţin 4 dintre cuvintele enumerate: *baltă* (v. ILR p. 328; de altfel, deşi DLRM şi DEX nu menţionează faptul, termenul are corespondent în alb. *baltë* „noroi, mocirlă"), *băiat, butuc, pînză* (pentru toate trei v. de ex. I.I. Russu *Limba traco-dacilor* p. 215—216). După Gh. Ivănescu (*Istoria limbii române* p. 259) „este foarte probabil" ca şi *hoţ* să aibă aceeaşi obîrşie. În afara acestor 5 termeni, calculul nostru ar mai indica deci existenţa a încă alte vreo 2 unităţi de substrat în grupul *moc, păcală, pungă, străin, tearfă, vorbă.*

Coborînd în zona productivităţilor mai mici (7—8 derivate) ne aşteptăm să mai întîlnim, potrivit tabelului 10.17.5., vreo 6 cuvinte autohtone printre următoarele 16: *bade, ban, brînză, burduf, doină, migăli, moţ, perie, prunc, scălîmb, stăpîn, sută, şovăi, ţipa, viscol, zburda.* Şi aici, pentru *brînză* (ILR p. 354), *burduf* (id. p. 329), *doină* (id. p. 355), *zburda* (id. p. 329), *prunc* şi *viscol* (ambele la I.I. Russu pag. 216) s-a sugerat provenienţa din substrat. Cum e vorba de exact 6 termeni, pare puţin probabil ca printre cei 10 rămaşi să mai fie şi alte cuvinte autohtone (în ciuda marii vechimi pe care o atestă, de pildă, trecerea în *î* a vocalei *á* urmată de *n* din *stăpîn*, schimbare fonetică anterioară despărţirii dialectelor româneşti).



Scanned with OKEN Scanner

Alte vreo 5 cuvinte de substrat ar fi de căutat printre cele 23 de unităţi de productivitate 5—6: *băga, buimac, burtă, colb, ghes, ghiont, gogoaşă, licări, mălai, mîhni, motan, paraginǎ, pisc, răbda, reteza, saia, ticală, ţugui, ţurcă, vuvă, zbînţui, zgudui.* În lista de cuvinte „sigur sau probabil" autohtone întocmită de I.I. Russu (op.cit. p. 215—216) figurează exact 5 dintre ele: *băga, burtă, ghes* (?), *ghiont* (?), *răbda* (semnele de întrebare aparţin autorului citat). Şi aici deci şansele de a mai descoperi alte elemente preromane sînt minime.

Cît priveşte grupul cuvintelor care au generat 3—4 derivate, situaţia este diferită. Conform datelor din tabelul 10.17.5., un număr de 27 de unităţi din totalul de 58, adică aproape 47%, pot proveni din substrat. Cum numai pentru 6 dintre ele *(baci, cîrlan, rezema, strugure, urdă, zer)* s-a propus originea autohtonă, rezultatul obţinut atrage atenţia asupra acestei zone a vocabularului ca posibilă rezervă de termeni traco-dacici încă nedepistaţi.

În schimb, în mulţimea cuvintelor de productivitate scăzută (1—2 derivate) probabilitatea apartenenţei la substrat este considerabil mai mică. De-abia vreo 9%, respectiv 19 unităţi din 202, ar putea descinde din etimoane autohtone. Această proveniență a fost deja propusă pentru: *bordei, buiestru* (?), *dezbăra* (?), *măceş, melc* (v. lista mai sus citată a lui I.I. Russu) şi *stînă* (ILR p. 356). Ar mai rămîne de căutat numai 13 cuvinte printre cele 196 rămase (6,6%).

### 10.18. Cuvinte de origine incertă

După cum s-a mai arătat, am grupat în această categorie, urmîndu-l pe D. Macrea, cuvintele a căror etimologie nu este complet elucidată, deşi o anumită vecinătate lexicală a lor a putut fi stabilită. Marca formală a apartenenţei la clasa cuvintelor de origine incertă este acel *comp.* (sau *cf.*) care precede în dicţionare termenul străin cu care pare a se înrudi cuvîntul în cauză. Este vorba, firește, de un inventar lexical foarte variat, asemănător prin neomogenitate cu cel reprezentat de domeniul cuvintelor cu etimologie necunoscută. În lista de mai jos, cuprinzînd cuvintele productive care au generat cel puţin 3 derivate, varietatea spectrului etimologic este pusă în lumină de vecinătăţile indicate în paranteză:

| | | |
|---|---|---|
| 1. | vai (cf. lat. vae) | 20 de derivate |
| 2. | pit(ic) (cf. v. sl. pitǫ) | 14 derivate |
| 3—4. | şold (cf. pol. szoldra), tăvăli, (cf. v. sl. valiti) | 12 derivate |



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 5—8. | frică (cf. gr. phrike), nene, (cf. bg. nenja, scr. nena), petic (cf. lat. pittacium), talpă (cf. magh. talp) | 11 derivate |
| 9—11. | flăcău (cf. v. sl. chlakŭ), tîlhar (cf. magh. tolvaj), ţeapă (cf. v. sl. čepati) | 10 derivate |
| 12—13. | cucoană (cf. ngr. kokkona), ursi (cf. ngr. orizo) | 9 derivate |
| 14—23. | broboadă (cf. bg. podbradka), bumbac (cf. scr. bumbak, bg. bubak), bură (cf. scr. bura), propti (cf. v. sl. podŭpŕti), slut (cf. ucr. slutyj), stropi (cf. v. sl. kropiti), ţuică (cf. scr. cujka), valmă (cf. rus. ucr. valom), zăpadă (cf. v. sl. zapadati), zimţ (cf. bg. zăbec, scr. zubac) | 8 derivate |
| 24—27. | ciocni (cf. bg. čukna, ucr. koknuti), glod (cf. magh. galád), preajmă (cf. v. sl. prěždi), scrijea (cf. scr. ucr.) | 7 derivate |
| 28—38. | alamă (cf. it. lama), bubă (cf. ucr. buba), clanţă (cf. onom. clanţ), dop (cf. sas Dop), mătase (cf. lat. metaxa), piţiga (cf. it. pizzicare), pripi (cf. bg. pripra), sas (cf. magh. szasz), smîntînă (cf. bg. ucr. smetana), trîndav (cf. scr. truntav), zdreanţă (cf. v. sl. sŭdranŭ) | 6 derivate |
| 39—57. | balmoş (cf. magh. bálmos), ban (cf. scr. ban, magh. bán), cicală (cf. it. cicala), ciot (cf. it. ciotto), ciupi (cf. bg. čupjat, scr. čupati), cîrlig (cf. bg. kărlik), coşcov (cf. rus kočkovatyj), găsi (cf. v. sl. gasiti), hîrb (cf. bg. hărbel), îmbia (prob, lat.* inviare), jindui (cf. v. sl. žendati), mălură (cf. ngr. meluri), mîl (cf. ucr. mul), mînji (cf. v. sl. mazati), posomorî (cf. magh. szomorú, rus. pásmuritsja), scamă (cf. lat. squama), sîmbră (cf. magh. cimbora), trînti (cf. bg. tărtja), zbîrci (cf. bg. bărča) | 5 derivate |
| 58—76. | ademeni (cf. magh. adomány), baie (cf. v. sl. banja), cofă (cf. bg. scr. kofa), copoi (cf. magh. kopo), laie (cf. pol. ucr. laja), povaţă (cf. pol. powodca), răzeş (cf. magh. részes), scheuna (prob. onom.), sculă (cf. | |



Scanned with OKEN Scanner

ngr. skula), sîc (cf. tc. sik) sprinten (cf. scr. spretan), steag (cf. rus. stjag), stejar (cf. bg. stežer), stropşi (cf. bg. strošvam), taler (cf. bg. taler), tîrşii (cf. magh. törzsölni), ului (cf. rus. ulovitь), vînj (cf. v. sl. vonžǔ), zvîrcoli (cf. bg. vărgaljam) 4 derivate

77—108. bădăran (cf. magh. badaró), bîigui (cf. magh. bolyo(n)gni), ciucure (cf. lat. cicculum), ciung (cf. it. cionco), ciunt (prob. contaminare ciot + ciung), clonţ (cf. bg. kljunec), cocli (cf. bg. kotljasvam), cocoloş (prob. onom.), felie (cf. ngr. feli), ferfeniţă (cf. ucr. ferfélic'a), hăţ (prob. onom.), huideo (cf. bg. scr. ujdo), iscăli (cf. v. sl. iskaljati), jumuli (cf. magh. gyomlálni), leică (cf. bg. lъžica), măslui (cf. rus. maslo), murui (cf. ucr. murovati), opaiţ (cf. v. sl. opajeci, rom. văpaie), opinti (cf. rus. opjatit'), pili (cf. ţig. pilo), rablă (cf. bg. hrăbla), ragilă (cf. germ. Raffel), răvăşi (cf. scr. rovašiti), scoroji (cf. v. sl. skora), spîrc (cf. ucr. špyrka, pol. szperka), şişcă (cf. scr. šiška), şontorog (cf. magh. dialectal czomtorag), tehui (cf. magh. tahonya), tulă (cf. scr. trlo), zăgaz (cf. scr. zagata, zagatiti), zălezi (cf. ucr. zaležaty, rus. zalezat'), zvîrli (cf. bg. hvărlja, scr. vrljiti) 3 derivate

Situaţia generală a acestei categorii de cuvinte este reflectată în tabelul 10.18. şi în figura 10.18.

*Tabelul 10.18*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 251 | 159 | 108 | 76 | 57 | 38 | 27 | 23 | 13 | 11 |
| ln (Nr. cuv.-bază) | 5,53 | 5,07 | 4,68 | 4,33 | 4,04 | 3,64 | 3,30 | 3,14 | 2,56 | 2,40 |

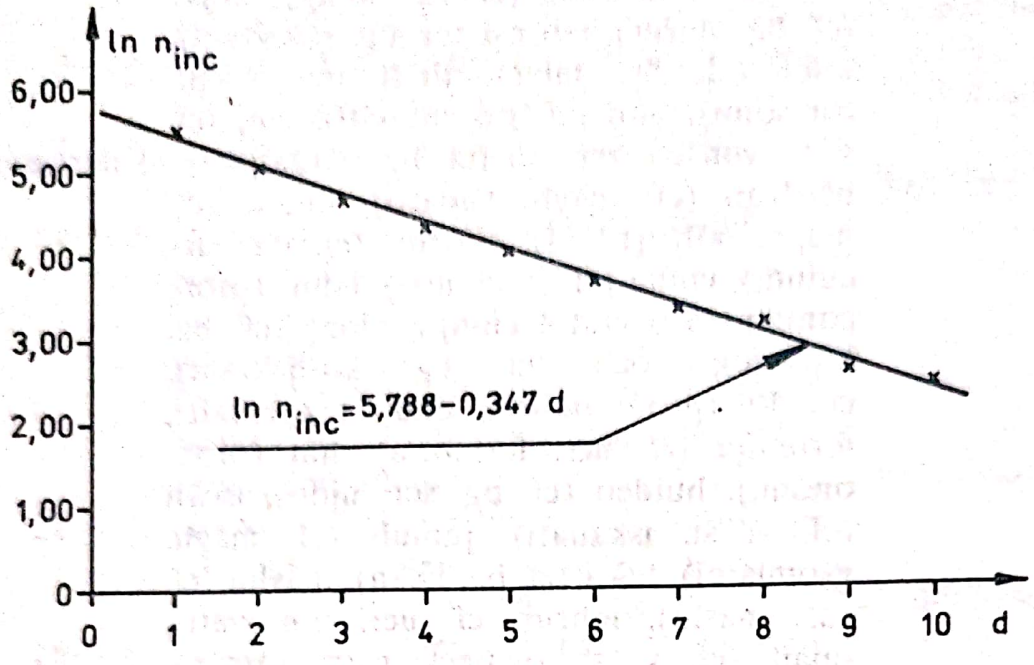
Expresia matematică a dreptei caracteristice este:

$$ln\, n_{inc} = 5{,}788 - 0{,}347\, d \qquad (10.18)$$



Scanned with OKEN Scanner

FIG. 10.18




Raportat la panta celorlalte drepte, determinate anterior, parametrul $m = 0{,}347$ indică apartenenţa precumpănitoare a cuvintelor de provenienţă incertă la straturile lexicale mai vechi (latin, vechi slav, bulgar, onomatopeic, maghiar), constatare ce concordă cu indicaţiile etimologice transcrise în paranteze. Este interesant de remarcat că termenii de origine nesigură se vădesc astfel mai vechi, în medie, decît cuvintele cu etimologie necunoscută, argument suplimentar împotriva exagerărilor daciste în interpretarea acestui din urmă material.

Să mai observăm că, în ciuda dificultăţii de a-l explica pe rom. *vai* prin lat. *vae* (pronunţat *ve* în vremea implantării limbii latine în Dacia), poziţia de lider detaşat a acestui cuvînt în clasamentul prezentat stă mărturie pentru marea sa vechime (să fi existat în substrat o interjecţie asemănătoare?)

## 10.19. Cuvinte formate în româneşte cu material lexical latin

În privinţa formaţiunilor lexicale româneşti, trebuie spus că ele aparţin mai multor categorii distincte. Prima, şi de departe cea mai vastă, este chiar cea a cuvintelor derivate care fac obiectul cercetării noastre. O a doua ar fi reprezentată de unităţile de provenienţă onomatopeică, generatoare la rîndul lor de derivate, de care ne-am



Scanned with OKEN Scanner

mai ocupat în paragraful 10.10. În fine, a treia este alcătuită din cuvintele compuse din material lexical moştenit sau împrumutat. Procedeul compunerii nefiind în româneşte un mijloc uzual de formare a unităţilor lexicale noi (cum se întîmplă pe scară largă în greacă sau germană), cuvintele compuse ocupă în vocabularul limbii române o poziţie marginală, cu o singură importantă excepţie: termenii alcătuiţi prin compunere din material lexical de provenienţă latină. Fondul principal lexical definit de Al. Graur a reţinut 30 de cuvinte „formate în româneşte din elemente latine": *adins, alde, alerga, amănunt, amîna, boci, căci, cerşi, dacă, deci, destul, deşi, din, greaţă, împreună, încît, îndată, leagăn, mănuşă, mulţumi, nebun, numai, parcă, peste, prin, printre, stărui, totuşi, unealtă, zbuciuma.* Chiar dacă o parte dintre ele sînt derivate[1] (*leagăn, mănuşă, nebun*), iar altele constituie termeni primari, nici compuşi, nici derivaţi (*alerga* ⟨ lat **allargare, greaţă* ⟨ lat **grevitia, stărui* ⟨ bg. *staraivse* (!)), rămîn încă destule cuvinte compuse, fapt ce atestă însemnătatea categoriei lexicale respective. Prin contrast, nici o unitate alcătuită prin compunere din material nelatin nu a ajuns pînă în FPL (cele două cuvinte „formate în româneşte din elemente slave" sînt simple derivate).

Discrepanţa dintre compusele de sursă latină şi cele formate din elemente de alte origini este subliniată pregnant şi de statistica întocmită pe baza DLRM. Ca număr total, ponderea acestor două clase este destul de apropiată (264 faţă de 203), dar cuvintelor formate cu material latin le datorăm 122 derivate, faţă de numai 33 cîte au generat celelalte. Slaba productivitate a acestora din urmă le face inapte pentru o tratare statistică, aşa încît ne vom rezuma la examinarea situaţiei celor dintîi. Compusele româneşti alcătuite din elemente de sorginte latină care s-au dovedit productive (la nivelul materialului cuprins în DLRM) sînt următoarele:

| | | |
|---|---|---|
| 1. | mulţumi | 10 derivate |
| 2. | împrejur | 9 derivate |
| 3. | departe | 8 derivate |
| 4—5. | destul, unealtă | 6 derivate |
| 6—7. | cuminte, împreună | 5 derivate |

[1] Distincţia dintre derivate şi compuse nu este întotdeauna uşor de făcut, căci există prefixe formal identice cu unele prepoziţii *(a, în)*, iar atunci cînd al doilea element alcătuitor e un substantiv nu mai ştim bine cum să îl interpretăm pe primul. Autorii DLRM apreciază, de exemplu, că particula *în* este prepoziţie în compusul *înlocui* ( < *în* + *loc*), dar prefix în *înstări*, opţiune, după noi, greu de argumentat.



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 8—9. | deplin, dumneata | 4 derivate |
| 10—13. | adins, alături, codalb, înlocui | 3 derivate |
| 14—25. | aminte, amîna, batjocoră, binecuvînta, ceva, cincizeci, cineva, dreptunghi, întroloca, laolaltă, peste, săvai | 2 derivate |
| 26—54. | alaltăieri, altădată, binefacere, binevoi, cincisprezece, cîtva, cumsecade, dalb, dis-dedimineaţă, doisprezece, douăzeci, întruni, laolaltă, nouăsprezece, nouăzeci, optsprezece, optzeci, paisprezece, patruzeci, sinucide, şaisprezece, şaizeci, şaptesprezece, şaptezeci, treisprezece, treizeci, unsprezece, untdelemn, vinars | 1 derivat |

*Tabelul10.19.1.*

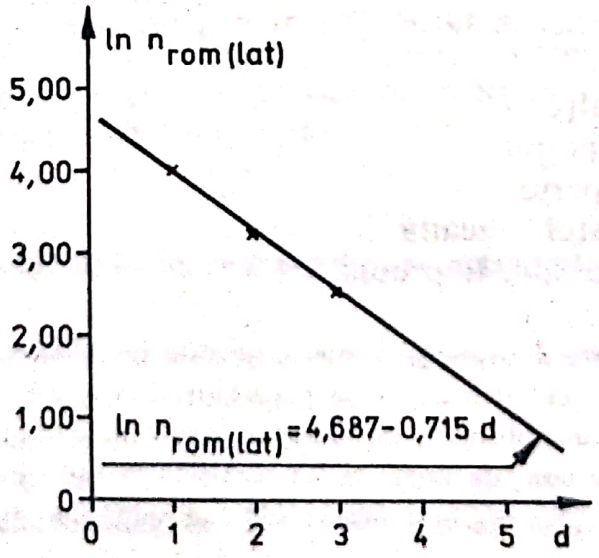
| NR. DERIVATE ≥ | 1 | 2 | 3 |
|---|---|---|---|
| Nr. cuv.-bază | 54 | 25 | 13 |
| ln (Nr. cuv.-bază) | 3,99 | 3,22 | 2,56 |

Dreapta caracteristică reprezentată în fig. 10.19.1. are ecuaţia:

$$ln\ n_{rom(lat)} = 4{,}687 - 0{,}715\ d \qquad (10.19.1)$$

De data aceasta, panta relativ mare (numai cu puţin inferioară celei aferente elementului latin savant) nu mai poate fi explicată prin caracterul recent al apariţiei în limbă a cuvintelor în cauză. Simpla

FIG. 10.19.1



parcurgere a listei de mai sus este suficientă pentru a ne convinge că unităţile ce o compun sînt, aproape toate, cu mult mai vechi decît o sugerează relaţia pantă-vechime aşa cum o cunoaştem din analiza situaţiilor anterioare. Neconcordanţa pare să provină din numărul prea mic de puncte (numai 31) avute la dispoziţie pentru trasarea dreptei caracteristice. În nici unul dintre cazurile deja studiate fondul de date de bază nu era atît de sărac, cu excepţia celui privitor la aportul lexical italian. Nepotrivirea constatată acum ne atrage atenţia asupra valorii strict informative a parametrilor calculaţi şi pentru influenţa italiană.

Se înţelege de ce ne oprim aici cu inventarierea componentelor etimologice ale vocabularului limbii române, lăsînd la o parte unele aporturi lexicale (săsesc, ţigănesc, englez) cu o prezenţă cantitativ prea modestă pentru a permite o tratare semnificativă cu metodele prezentate mai sus.

Pentru a ne convinge că un spor de date de bază ar putea schimba considerabil forma ecuaţiei (10.19.1), şi, implicit, concluziile ce decurg din ea, să facem următoarea experienţă:

Am convenit, după cum se ştie, să nu luăm în considerare, la trasarea dreptelor caracteristice, decît clasele de cuvinte productive ce conţin minimum 10 elemente, socotind că prin coborîrea la cifre mai mici s-ar vicia rezultatele unui calcul care se întemeiază în principal pe date de natură statistică. Să ne revizuim acum, cu titlu experimental, poziţia, reducînd pragul de exigenţă pînă la a lua în considerare toate clasele lexicale care posedă cel puţin 5 membri. Ca urmare, tabelul 10.19.1. se mai îmbogăţeşte cu încă 3 coloane (v. tabelul 10.19.2).

*Tabelul 10.19.2.*

| NR. DERIVATE ≥ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 54 | 25 | 13 | 9 | 7 | 5 |
| ln (Nr. cuv.-bază) | 3,99 | 3,22 | 2,56 | 2,20 | 1,95 | 1,61 |

Tripleta de puncte nou apărute, de abscise $d = 4, 5, 6$, (v. fig. 10.19.2.) modifică sensibil înclinarea dreptei obligate să treacă printre ele. Ecuaţia (10.19.1) capătă acum forma:

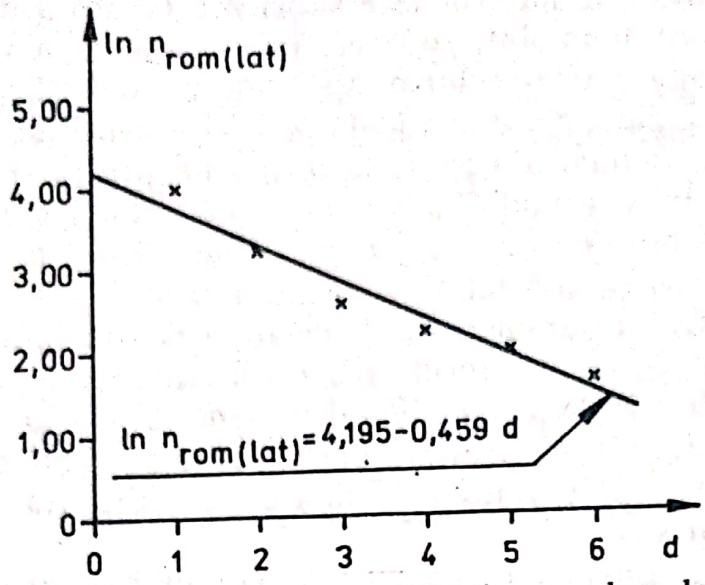
$$ln\, n_{rom\ (lat)} = 4{,}195 - 0{,}459\, d \qquad (10.19.2.)$$

Panta mult mai lină decît cea determinată iniţial ($m = 0{,}459$ în loc de 0,715) aruncă dintr-odată cu secole în urmă epoca de formare a cuvintelor compuse din material lexical de origine latină. Că o atare



Scanned with OKEN Scanner

FIG. 10.19.2




ipoteză este mai plauzibilă decît cea dintîi ne-o dovedeşte şi calculul numărului scontat de derivate al celui mai productiv cuvînt: ecuaţia (10.19.1) furnizează $d_{max} = 4{,}687 : 0{,}715 = 6{,}55$, iar ecuaţia (10.19.2.) $d_{max} = 4{,}195 : 0{,}459 = 9{,}14$, această din urmă valoare fiind mult mai apropiată de productivitatea reală ($d = 10$) a verbului *mulţumi* care deţine poziţia fruntaşă în lista termenilor-bază. Sîntem datori să observăm că, în cazul unui fond de date mai bogat, o extindere similară a şirului cu care se operează nu ar fi fost de natură să afecteze la fel de mult poziţia dreptelor caracteristice[1]. E drept că, dacă adăugarea unui nou punct la mulţimea celor 22 luate în considerare în calculul elementului latin nu antrenează schimbări sesizabile în aspectul dreptei caracteristice, în schimb atunci cînd se dispune doar de 5 puncte, ca în cazul influenţei germane, turceşti, ruseşti, ucrainiene sau latine savante, pot apărea (nu în mod necesar) unele modificări. Situaţia nu poate fi evitată şi, în orice caz, coborîrea sau ridicarea pragului, e drept convenţional, al celor 10 cuvinte nu asigură o rezolvare mai bună a problemei. În general, aşa cum o dovedeşte experienţa multor încercări pe care le-a făcut anterior autorul acestor rînduri, adoptarea unui prag mai ridicat reduce prea mult numărul de puncte utilizabile, mărind arbitrariul trasării dreptei, în vreme ce scăderea sub 10 a limitei alese potenţează în mod nepermis importanţa unor cuvinte foarte productive care pot constitui simple excepţii (cum sînt, de exemplu, în cazul elementului turcesc, cele cîteva elemente de origine cumană).

[1] Lăsăm cititorului ca exerciţiu verificarea acestei afirmaţii pentru componentele etimologice deja studiate ale lexicului limbii române.



Scanned with OKEN Scanner

# 11. Implicaţii, virtuţi şi limite ale modelului matematic

Trecerea în revistă a principalelor 19 componente etimologice ale vocabularului românesc, care acoperă cca 97% din numărul total de cuvinte productive, ne-a prilejuit constatarea existenţei unei corelaţii cu caracter de lege lingvistică, valabilă pentru fiecare aport lexical omogen luat în parte. Dacă se notează cu $d$ numărul de derivate ale unui cuvînt şi cu $n$ ($\geq d$) numărul de cuvinte-bază care au generat cel puţin $d$ derivate, întregul material lexical de o aceeaşi origine se supune unei ecuaţii de forma:

$$n\,(\geq d) = exp(a - md) \text{ unde } a,\ m = \text{constant} \qquad (11.1)$$

Să ne oprim puţin asupra semnificaţiei lingvistice a relaţiei de mai sus.

Vom începe prin a observa că numărul $n(d)$ al cuvintelor care posedă exact $d$ derivate poate fi lesne calculat pornind de la ecuaţia (11.1) şi că el are expresia:

$$n(d) = n\,(\geq d) - n(\geq (d+1)) \cdot = exp\,(a - md) - exp\,(a - m(d+1))$$

sau

$$n(d) = exp\,(a - md)\,(1 - exp\,(-m)) \qquad (11.2)$$

Evident, numărul de cuvinte de productivitate $(d+1)$ se obţine înlocuind în (11.2) argumentul $d$ prin $(d+1)$:

$$n(d+1) = exp\,(a - m(d+1))\,(1 - exp\,(-m))$$

$$n(d+1) = exp\,(a - md)\,exp\,(-m)\,(1 - exp\,(-m)) \qquad (11.3)$$

Împărţind acum (11.2) la (11.3), expresiile $exp\,(a - md)$ şi $(1 - exp\,(-m))$ aflate atît la numărător cît şi la numitor se simplifică şi rezultă relaţia simplă:

$$\frac{n(d)}{n(d+1)} = exp\,m \qquad (11.4)$$

valabilă pentru orice $d$.

Rezultatul (11.4) are o semnificaţie deosebită. Potrivit relaţiei (11.4), pentru o componentă etimologică dată, numărul de cuvinte-bază care au dat fiecare naştere la cîte $d$ derivate este de $exp\,m$ ori



Scanned with OKEN Scanner

mai mare decît al celor care au produs exact cu un derivat mai puţin. Astfel, de exemplu, pentru fondul latin moştenit, cuvintele ce posedă un singur derivat sînt de *exp* 0,230 = 1,26 de ori mai multe decît cele care au două derivate. Acestea din urmă depăşesc tot de 1,26 ori numărul cuvintelor cu trei derivate, care sînt, la rîndul lor, de 1,26 ori mai numeroase decît cele cu patru derivate şi aşa mai departe. Pentru elementul „francez" lucrurile se prezintă la fel, cu singura deosebire că raportul este de data aceasta de *exp* 1,092 = 2,98 ≈ 3. Înseamnă că pentru orice cuvînt de origine „franceză" care a ajuns să aibă, să spunem, 5 derivate trebuie să existe circa 3 cuvinte cu cîte 4 derivate, 3 × 3 = 9 cuvinte cu 3 derivate, 9 × 3 = 27 de cuvinte cu 2 derivate şi 27 × 3 = 81 de cuvinte cu un singur derivat. În schimb, existenţa unui cuvînt moştenit din latineşte care să fi produs 5 derivate, presupune prezenţa în vocabularul limbii române a numai $1{,}26^4 = 2{,}5$ cuvinte cu un singur derivat. Pătrunderea în elita unităţilor lexicale foarte productive este aşadar mult mai dificilă pentru un cuvînt „francez" decît pentru unul latinesc. Din această pricină, performanţa, destul de obişnuită în cazul cuvintelor de origine latină, de a poseda cîte 10 derivate (condiţie îndeplinită de peste 40 de cuvinte-bază) este de neatins pentru împrumuturile „franceze", apariţia unui singur cuvînt atît de productiv necesitînd o rezervă pasivă de 18 533 de cuvinte de origine „franceză" care să fi dat naştere cîte unui derivat şi, în general, de 27 82 de unităţi „franceze" productive. Dacă ne mai îndoiam de legitimitatea excluderii dintre împrumuturile franceze a verbului *semna* avem acum dovada formală că acest cuvînt nu poate fi un neologism. Conform legii ce guvernează distribuţia după productivitate a cuvintelor „franceze", un atare cuvînt nu ar fi putut ajunge să posede 18 derivate decît în ipoteza că vocabularul limbii române ar fi conţinut nu mai puţin de 172 612 850 de unităţi lexicale productive de provenienţă „franceză"!

Pentru a atinge aşadar un anumit nivel de productivitate, împrumutul trebuie să dispună de o platformă lexicală corespunzătoare, să se sprijine pe un număr de cuvinte de aceeaşi origine, cu atît mai mare cu cît panta dreptei caracteristice este mai mare. Or, s-a observat că panta depinde direct de vechimea influenţei lexicale respective. Un loc fruntaş în ierarhia unităţilor productive se obţine deci prin efectul combinat al vechimii şi al abundenţei de cuvinte de aceeaşi origine, nici una dintre aceste două condiţii nefiind singură îndestulătoare. Astfel, în ciuda masei lor enorme, neologismele nu au reuşit să acceadă la ranguri superioare pe scara productivităţii din pricina vechimii lor insuficiente. Tot atît de puţine sînt însă şi şansele ca un cuvînt vechi, dar izolat, provenind dintr-o limbă care nu



Scanned with OKEN Scanner

a mai transmis românei alte elemente să parvină la o poziţie fruntaşă. Posibilele împrumuturi din limbile vechi germanice (gotă, gepidă) sau din avară, despre existenţa cărora s-a discutat mult în lingvistica noastră, aparţin în consecinţă, dacă există, în pofida vîrstei lor respectabile, categoriei unităţilor ne- sau slab productive. Conchidem că generarea de derivate nu este numai o problemă individuală a cuvîntului generator, ci ea priveşte şi depinde de totalitatea materialului lexical de aceeaşi obîrşie pătruns în limba română şi probabil de ansamblul interacţiunilor lingvistice dintre cele două idiomuri implicate. Funcţionarea vocabularului ca sistem ni se dezvăluie şi în această observaţie.

Cîteva explicaţii sînt necesare în legătură cu modul cum au fost obţinute rezultatele de pînă aici. După cum s-a putut constata, demersul nostru a cunoscut două etape distincte. Într-o primă etapă s-au efectuat inventarierea şi clasificarea după criteriile propuse (origine, productivitate) a materialului lexical ce făcea obiectul analizei (totalitatea cuvintelor din DLRM), pentru ca, pe această bază, să se construiască în etapa următoare un număr de 19 ecuaţii menite să dea seamă de comportarea sub aspectul productivităţii lexicale a principalelor 19 componente etimologice ale limbii române. Împreună, cele 19 ecuaţii constituie un *model matematic*, unul dintre multele posibile ale vocabularului. Odată construit, *acest model* a devenit obiectul nostru de studiu. În măsura în care se referă la model, iar nu la realitatea pe care el o reflectă, toate rezultatele şi concluziile ce decurg din acestea prezintă un anumit grad de aproximaţie. Afirmam, de exemplu, că pentru elementul latin raportul dintre numărul cuvintelor de productivitate $d$ şi numărul celor de productivitate $(d + 1)$ este de 1,26, oricare ar fi $d$. Aserţiunea este exactă în ceea ce priveşte modelul matematic, dar, raportată la situaţia concretă a fondului lexical latin, ea rămîne aproximativă. Pornind de la bilanţul cuprins în tabelul 10.1., se pot lesne calcula valorile reale ale raportului respectiv pentru fiecare productivitate $d$ distinctă. Tabelul 11.1. sistematizează aceste valori.

*Tabelul 11.1*

| $d$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{n(d)}{n(d+1)}$ | 1,80 | 1,16 | 1,05 | 1,26 | 0,89 | 1,.47 | 1,59 | 1,30 | 0,96 | 1,50 | 2,14 |

| $d$ | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{n(d)}{n(d+1)}$ | 0,82 | 1,31 | 1,08 | 1,50 | 1,33 | 1,20 | 1,00 | 1,00 | 1,25 | 1,00 | 1,00 |



Scanned with OKEN Scanner

Şirul astfel obţinut este alcătuit, după cum se vede, din valori destul de diferite, dar acestea oscilează, ca în model, în jurul cifrei de 1,26. Să notăm că nici nu ar fi fost practic posibil ca, în condiţiile în care avem de a face numai cu numere întregi, raportul lor să fie peste tot egal cu 1,26 (ecuaţiile modelului, în schimb, operează şi cu numere neîntregi de cuvinte).

Abaterile, destul de mari, de la valoarea medie ale cifrelor din tabelul 11.1. sînt considerabil diminuate dacă se ia în considerare şirul valorilor integrate (adică în loc de numărul cuvintelor cu $d$ derivate ne raportăm la numărul cuvintelor cu *cel puţin* $d$ derivate). Era şi firesc să se întîmple aşa, întrucît tocmai acest al doilea şir a constituit punctul de plecare pentru calculul parametrilor dreptei caracteristice. Tabelul 11.2. evidenţiază gruparea, mult mai restrînsă, a rapoartelor dintre mulţimile succesive de cuvinte în jurul valorii medii, care este şi aici tot[1] 1,26.

*Tabelul 11.2.*

| $d$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{n(\geq d)}{n(\geq d+1)}$ | 1,35 | 1,20 | 1,20 | 1,24 | 1,23 | 1,35 | 1,31 | 1,24 | 1,23 | 1,32 | 1,27 |

| $d$ | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\frac{n(\geq d)}{n(\geq d+1)}$ | 1,28 | 1,24 | 1,23 | 1,27 | 1,22 | 1,19 | 1,19 | 1,24 | 1,24 | 1,31 | 1,44 |

Cu excepţia ultimului număr, ecarturile în raport cu media nu depăşesc aici 7% (în cazul precedent urcau pînă la 70%!), dovadă că ecuaţiile stabilite pe valori integrate modelează fenomenul lingvistic cu o precizie absolut mulţumitoare. Nu e mai puţin adevărat că ele nu fac decît să aproximeze realitatea modelată, căci modelul reflectă

---

[1] La întrebarea, legitimă, de ce panta medie rămîne identică, indiferent dacă se referă la şirul simplu sau la cel integrat, răspunsul îl dau proprietăţile matematice ale funcţiei exponenţiale $f(x) = exp\, x$, pe care operaţiile de integrare şi derivare o lasă neschimbată: $\int exp\, x\, dx = exp\, x$; $(exp\, x)' = exp\, x$. Ca urmare, subzistă $(exp\,(a - mx))' = -m\, exp\,(a - mx)$ şi $\int exp\,(a - mx)\, dx = -\frac{1}{m}\, exp\,(a - mx)$ şi invarianţa pantei este demonstrată.



Scanned with OKEN Scanner

tendinţele generale, dar nu îşi propune să dea seamă de cazurile individuale. Tocmai în această slăbiciune rezidă şi forţa sa. Nu putem explica cu ajutorul ecuaţiilor stabilite comportarea concretă a unui anumit cuvînt (în treacăt fie zis, lingvistica dispune de suficiente metode de cercetare a statutului şi evoluţiei unităţilor lexicale particulare), însă putem obţine o imagine globală a structurii şi dinamicii vocabularului, pe care cercetarea atomistică de tip tradiţional nu era în măsură să ne-o ofere. Sacrificînd individualul, modelul generalizează observaţiile privind productivitatea cuvintelor de diferite origini pînă la a caracteriza aportul fiecărei limbi ce a participat la formarea lexicului românei moderne printr-un simplu cuplu de valori $\langle a, m \rangle$, reprezentînd cele două constante din ecuaţia dreptei caracteristice: $ln\, n = a - md$. Fiecărei surse etimologice a vocabularului românesc îi corespunde un cuplu $\langle a, m \rangle$ distinct şi, prin aceasta, comportarea sub aspectul productivităţii a cuvintelor de o anumită provenienţă devine reprezentabilă sub forma unui unic punct din planul de coordonate $(a, m)$. Un nor de circa 20 de astfel de puncte caracterizează în totalitate distribuţia pe origini şi potenţial de generare de cuvinte noi a ansamblului lexicului limbii române. O treaptă de generalizare mai înaltă nu pare posibilă şi nici nu ar fi de dorit, avînd în vedere pericolul ca modelul, devenit prea „încăpător", să-şi piardă relevanţa, puterea explicativă.

Este important de subliniat că, în ciuda caracterului său sincronic (el reflectă în mod strict situaţia limbii române de la jumătatea secolului al XX-lea), modelul matematic oferă multe sugestii de natură diacronică, dat fiind că, într-un fel sau în altul, ambii membri ai cuplului $\langle a, m \rangle$ depind de timp. Parametrul $a$ reprezintă logaritmul natural al numărului total de cuvinte de o origine dată pe care tinde să le mai conserve limba vie la mult timp după ce influenţa respectivă a încetat să se exercite.

În cazul fondului latin moştenit, diferenţa dintre $a = 7{,}475$ şi logaritmul numărului de cuvinte de sorginte latină cuprinse în dicţionar ($= 7{,}50$) este aproape neînsemnată. Prin contrast, pentru neologismele grupate sub eticheta „francez", valoarea $ln\, n_{total} = 9{,}63$ depăşeşte copios constanta $a = 8{,}70$ din ecuaţia dreptei caracteristice[1]. O dependenţă şi mai clară de timp prezintă parametrul $m$. Funcţia $m(t)$ necesită o analiză mai detaliată pe care o întreprindem în paragraful următor.

[1] Decalajul dintre valorile logaritmate nu pare spectaculos, dar, în termeni de număr de cuvinte, el reprezintă o diferenţă copleşitoare: 15.000 faţă de numai 6.000.



Scanned with OKEN Scanner

## 12. Datarea împrumuturilor

Dificultatea elaborării unei formule de calcul care să exprime relaţia dintre vechime şi panta dreptei caracteristice nu este în primul rînd de natură matematică. Chiar dimpotrivă. Forma relaţiei e destul de limpede sugerată de modul cum variază valoarea parametrului $m$ pentru diferitele aporturi lexicale mai vechi sau mai noi. Începînd de la 0,230 pentru fondul latin moştenit (ceea ce corespunde unui unghi ceva mai mic de 15°), $m$ creşte pînă la 1,092 în cazul neologismelor zise „franceze" (puţin peste 47°). Prin urmare, cu cît materialul lexical este mai recent, cu atît panta devine mai abruptă. Între vechimea medie $t_m$ şi panta $m$ subzistă o relaţie de proporţionalitate inversă de tipul: $t_m \sim \frac{1}{m}$. Să presupunem că:

$$(t_m)^\alpha = c\left(\frac{1}{m}\right)^\beta \quad \text{unde } c \text{ este o constantă} \tag{12.1}$$

Se observă că pentru $t_m = 0$, avem $m \to \infty$, ceea ce corespunde realităţii, întrucît în chiar momentul pătrunderii lor în limbă cuvintele nu posedă încă derivate pe teren românesc, fapt ce se traduce matematic prin orientarea verticală a dreptei caracteristice (tangenta unghiului drept este, se ştie, infinită). Extrăgînd rădăcina de ordinul $\alpha$ din ambii membri, se obţine:

$$t_m = km^{-\gamma} \tag{12.2}$$

unde $\gamma = \beta/\alpha$, iar $k$ e o altă constantă, egală cu $\sqrt[\alpha]{c}$

Ecuaţia (12.2) nu este singura care satisface condiţiile date, dar e cea mai simplă dintre toate şi ni se pare normal ca încercările de a găsi forma optimă a relaţiei vizate să înceapă cu ea. Pe baza cunoaşterii a două cupluri particulare $\langle t_m, m \rangle$ se pot determina valorile constantelor $\gamma$ şi $k$.

Abia aici încep dificultăţile. Cunoaştem oare cu exactitate vechimea medie a cuvintelor de o origine determinată? Evident că nu, dar pentru unele aporturi se poate face o estimare ceva mai sigură decît pentru altele. Astfel, faţă de anul 1950, care este punctul nostru de referinţă, vechimea medie a cuvintelor latine moştenite este probabil cuprinsă între 1700 şi 1800 de ani. Raportate la acelaşi moment, cuvintele



Scanned with OKEN Scanner

„franceze" pot data în medie de 70—80 de ani. Să vedem la ce rezultate ne conduce acceptarea unei atari supoziţii.

Aplicînd, pe rînd, formula (12.2) elementului latin şi celui „francez", vom avea:

$$t_{m\,lat} = k(m_{lat})^{-\gamma} \tag{12.3}$$

$$t_{m\,fr} = k(m_{fr})^{-\gamma} \tag{12.4}$$

Raportul expresiilor (12.3) şi (12.4) este:

$$\frac{t_{m\,lat}}{t_{m\,fr}} = \left(\frac{m_{fr}}{m_{lat}}\right)^{\gamma} = \left(\frac{1{,}092}{0{,}230}\right)^{\gamma} = 4{,}748^{\gamma} \tag{12.5}$$

Potrivit ipotezelor de mai sus, raportul $\frac{t_{m\,lat}}{t_{m\,fr}}$ variază între 21,25 (în situaţia cînd $t_{m\,lat} = 1.700$ şi $t_{m\,fr} = 80$) şi 25,71 (în cazul $t_{m\,lat} = 1.800$ şi $t_{m\,fr} = 70$).

Din relaţia (12.5), acestor limite le corespund valorile $\gamma = 1{,}955 \div 2{,}076$. Aşadar $\gamma \approx 2$ şi relaţia (12.2) capătă forma comodă:

$$t_m = \frac{k}{m^2}$$

Pentru cuplul de date $t_{m\,lat} = 1700 \div 1800$, $m_{lat} = 0{,}230$, găsim $k \approx 90 \div 95$ şi deci:

$$t_m = \frac{90 \div 95}{m^2} \tag{12.6}$$

Să examinăm acum, cu un minimum de comentarii, datările ce rezultă din relaţia (12.6) pentru influenţele lexicale cărora li s-au determinat, în capitolul 10., ecuaţiile dreptelor caracteristice.

### 1. Aportul vechi slav

Cu panta $m_{sl} = 0{,}312$ (v. § 10.3.) se calculează că vechimea medie a împrumuturilor vechi slave este $t_{m\,sl} = 925 \div 975$, ceea ce corespunde perioadei anilor $975 \div 1025$. Rezultatul poate fi coroborat cu datele istorice. Într-adevăr, intervalul calculat coincide



Scanned with OKEN Scanner

nu numai cu centrul de greutate al extensiunii temporale extreme a înrîuririi slave (sec. VI—XVI), ci şi cu cel al perioadei mai restrînse de contact lingvistic efectiv, căci se ştie, pe de o parte, că „existenţa împrumuturilor slave înainte de secolele al VIII-lea şi al IX-lea nu poate fi probată cu argumente lingvistice" (ILR II, p. 373)[1], şi pe de altă parte, că „epoca de interferenţe a durat pînă prin sec. al XII-lea — XIII-lea"[2]. În plus, secolele X-XI reprezintă epoca formaţiunilor politice slavo-române atestate documentar şi totodată vremea adoptării liturghiei slave de către români[3]. Toate aceste elemente sînt de natură să confirme datarea obţinută cu ajutorul formulei (12.6) validînd modelul matematic propus[4]. Cu atît mai interesante vor fi, în consecinţă, rezultatele obţinute pentru influenţele în legătură cu care informaţiile istorice sînt lacunare sau lipsesc cu desăvîrşire.

## 2. Aportul bulgăresc

Dacă se ţine seamă şi de cuvintele cu etimologie mixtă bulgaro-sîrbocroată, panta dreptei caracteristice este $m_{bg} = 0{,}362$ (în caz contrar, $m_{bg} = 0{,}384$). Introdusă în relaţia (12.6), ea furnizează $t_{m\,bg} = 687 \div 725$, adică o vîrstă medie corespunzătoare perioadei $1225 \div 1265$. Este tocmai epoca acelei ascensiuni încă puţin cunoscute care a adus elementul bulgar la conducerea imperiului pînă atunci binaţional al vlahilor şi bulgarilor întemeiat de Asăneşti după răscoala antibizantină din 1186. Oricît de tentante ar fi speculaţiile ce se pot face pe marginea acestei coincidenţe, nu cedăm ispitei, cu atît mai mult cu cît ştim (v. paragraful 10.3) că o bună parte din materialul lexical considerat în calculul pantei drept avînd origine mixtă bulgaro-sîrbocroată este în realitate vechi slav. Includerea sa printre elementele bulgăreşti a împins centrul de greutate al acestei influenţe

---

[1] Demonstraţia, concludentă, a lui I. Pătruţ se bazează pe prezenţa în împrumuturile din slavă a vocalei *o* dezvoltată din *ă* (*a* scurt) după sec. al VIII-lea.

[2] *Istoria limbii române* coord.: FL. Dumitrescu, Ed. Didactică şi Pedagogică, 1978, p. 90.

[3] Ocuparea Transilvaniei de către unguri constituie un termen *ante quem* al acestei adoptări, deoarece procesul nu ar mai fi putut avea loc sub stăpînirea coroanei „apostolice".

[4] Să mai adăugăm celor spuse şi părerea lui Gh. Ivănescu după care „se poate admite că epoca de masive împrumuturi slave a avut loc după anii 800 sau 900 pînă prin secolul al XI-lea (*Istoria Limbii Române*, Ed. Junimea, Iaşi, 1980, p. 180).



Scanned with OKEN Scanner

către o fază mai timpurie, denaturînd întrucîtva perspectiva istorică asupra raporturilor lingvistice româno-bulgare, relaţii desfăşurate de-a lungul multor secole, pînă în imediata actualitate. Chiar dacă perioada convieţuirii celor două popoare într-un stat „al lor" va fi constituit un prilej de activizare a schimburilor lingvistice reciproce (inclusiv în hinterlandul transdanubian), este dificil de acceptat că afluxul de cuvinte bulgăreşti în limba română înainte de acest moment va fi fost atît de intens încît să compenseze în calculul vechimii medii vîrsta relativ scăzută a împrumuturilor mai noi. De aceea ni se pare cu mult mai verosimilă datarea obţinută cu ajutorul celeilalte valori a pantei, $m = 0{,}384$, obţinute pe baza exclusivă a materialului lexical de origine sigur şi unic bulgară. Vîrsta medie indicată de formula (12.6) este, în acest caz, $t_{m\,bg} = 610 \div 644$, ceea ce corespunde intervalului 1305 ÷ 1340, adică perioadei de constituire a voievodatului Ţării Româneşti, stat legat prin multe fire de ţaratele sud-dunărene şi care avea să ofere timp de secole azil valurilor succesive de imigranţi bulgari, refugiaţi din pricina persecuţiilor la care erau supuşi de autoritatea otomană[1].

### 3. Formaţiunile onomatopeice

Ne referim la acestea în succesiunea imediată a aportului bulgăresc, dată fiind identitatea pantei celor două drepte caracteristice: $m_{onom} = m_{bg} = 0{,}362$. Vechimea medie a ambelor componente ale vocabularului este deci aproximativ aceeaşi (a doua jumătate a sec. al XIII-lea). Această vîrstă apreciabilă îndreptăţeşte supoziţia că măcar cîteva dintre onomatopee datează încă din perioada românei comune. Căutarea lor constituie o problemă deschisă şi deloc simplă, avînd în vedere că prezenţa acestor onomatopee în dialecte diferite nu probează cu certitudine originea comună: caracterul lor imitativ lasă aproape întotdeauna loc şi pentru ipoteza genezei independente.

---

[1] Cronica lui Wavrin scrie bunăoară, referitor la evenimentele din timpul lui Vlad Dracul, că, în 1445, „Creştinii care treceau prin Bulgaria... s-au dus la Domnul Ţării Româneşti şi l-au rugat să le dea voie să treacă Dunărea şi să le dea sau să le lărgească loc în ţara lui". Proporţiile colonizării sînt indicate ceva mai încolo: „Domnul i-a primit pe creştinii bulgari, căci erau 12.000 de oameni, bărbaţi, femei şi copii, cu mărfuri şi vite" (apud Gh. I. Brătianu *Tradiţia istorică despre întemeierea statelor româneşti*, Ed. Eminescu, 1980, p. 72).



Scanned with OKEN Scanner

## 4. Elementul maghiar

Apropiat ca valoare de precedentele, parametrul $m_{magh} = 0{,}379$ indică, potrivit formulei (12.6), o vechime medie de 625 ÷ 660 de ani, fixînd centrul de greutate al influenţei maghiare în perioada 1290 ÷ 1325. Acest moment coincide cu frontiera propusă de Gh. Ivănescu pentru delimitarea celor două epoci distincte ale influenţei maghiare, evidenţiate, încă din secolul trecut, de Simion Mîndrescu *(Elemente ungureşti în limba română,* Bucureşti, 1892), perioada veche, anterioară întemeierii principatelor extracarpatice (sec. X—XIII), cînd înrîurirea maghiară s-a exercitat asupra limbii tuturor românilor, şi perioada mai nouă, caracterizată prin restrîngerea acestei influenţe la graiurile din regiunile direct supuse stăpînirii ungureşti (sec. XIV—XIX). După A. Rosetti, „primele elemente de limbă maghiară au pătruns în limba română în secolele al XI-lea — al XII-lea" (I.L.R. V, ed. V, 1964, p. 111). Ele sînt atestate documentar în textele slavo- române începînd cu sfîrşitul veacului al XIV-lea *(hotar* în 1392, *nemeş* în 1404)[1]. Lista, nici pe departe completă[2], a elementelor maghiare atestate în textele româneşti din secolul al XVI-lea prezentată de A. Rosetti (ILR V ed. V, 1964, p. 113—116) este suficientă pentru a da măsura extensiunii şi a adîncimii influenţei maghiare în faza sa mai veche. Semnificativ e faptul că, exceptîndu-i pe regionalul *cătană* şi pe discutabilul *şind(r)ilă* (probabil ⟨ germ. Schindel), toţi termenii proveniţi din ungureşte care au dat naştere la mai mult de 6 derivate *(şoim, gînd, socoti, ham, hotar, chip, locui, meşter, oraş, seamă, viclean, chibzui, tăgădui)* posedă atestări anterioare anului 1600. Dintre ei numai *han* şi *locui*[3] (acesta din urmă simţit probabil ca aparţinînd familiei cuvîntului de origine latină *loc*) lipsesc din fondul principal lexical propus de Al. Graur.

---

[1] ILR coord.: Fl. Dumitrescu, 1978, p. 100

[2] v. în această privinţă nota 1 de la p. 65.

[3] În legătură cu verbele terminate în *-ui* trebuie semnalată o inadvertenţă în clasificarea curentă pe origini a cuvintelor. Se recunoaşte că atît aceste verbe cît şi foarte productivul substantiv *gînd* au pătruns în româneşte pe filieră sud-slavă (A. Rosetti, ILR V, 1964, p. 112). Există însă un consens în a socoti drept limbă de origine a unui cuvînt românesc idiomul din care acesta ne-a parvenit în mod direct şi în acest sens elemente ca *bănui, bîntui, chibzui, făgădui* etc. nu pot fi considerate ca provenind din ungureşte, la fel cum *hîrtie* sau *drum*, odată trecute printr-un intermediar slav, nu mai au ce căuta în lista împrumuturilor greceşti.



Scanned with OKEN Scanner

S-a arătat într-un capitol anterior (v. paragraful 10.7) că luarea în considerare, alături de cuvintele de origine indubitabil maghiară, şi a termenilor de provenienţă nesigură ce posedă un corespondent unguresc la care dicţionarul trimite cu *cf.* nu modifică esenţial panta dreptei caracteristice. Rămîne să vedem dacă ecartul dintre cele două valori ale pantei nu se traduce totuşi printr-un decalaj temporal cît de cît însemnat. Introdusă în formula (12.6), panta $m = 0{,}387$ a dreptei (10.7.2) furnizează $t_m = 601 \div 634$. În noile condiţii, centrul de greutate al influenţei maghiare s-a deplasat aşadar din zona anilor $1290 \div 1325$ în intervalul $1315 \div 1350$, ceea ce revine la un decalaj de un sfert de veac pentru o schimbare de pantă de numai 2%. Acest rezultat ne atrage atenţia asupra marii sensibilităţi a formulei (12.6) care, după cum se observă, reacţionează pertinent la cea mai neînsemnată modificare a înclinării dreptei caracteristice. În acelaşi timp el ne avertizează să nu acordăm un credit exagerat periodizării influenţelor cu ajutorul modelării matematice, deoarece s-a constatat (v. de exemplu paragraful 10.13) că, cel puţin în unele cazuri, adăugarea sau omiterea numai a cîtorva cuvinte foarte productive se poate repercuta spectaculos asupra pantei şi, implicit (cu atît mai mult), asupra vechimii calculate.

### 5. Influenţa grecească

Panta $m = 0{,}501$ a dreptei (10.8), corespunde, potrivit relaţiei (12.6) unei vîrste medii de $359 \div 378$ de ani, adică perioadei $1570 \div 1590$. Ar fi inutil să căutăm vreo interpretare istorică acestui rezultat, ştiut fiind faptul că intervalul respectiv nu înseamnă ceva prin el însuşi, ci constituie numai efectul unei medieri ponderate a bimilenarei înrîuriri lexicale greceşti asupra limbii noastre. El arată cel mult că aportul mediogrecesc, bizantin, are o pondere îndeajuns de importantă pentru a contrabalansa masa considerabilă de cuvinte neogreceşti pătrunse în româneşte în perioada domniilor fanariote.

### 6. Împrumuturile germane

Epoca relativ timpurie (anii 1670—1685) pe care formula (12.6) o indică drept centru de greutate al influenţei germane se poate explica prin apartenenţa unei cote părţi destul de însemnate din materialul



Scanned with OKEN Scanner

lexical căruia DLRM îi atribuie origine germană, la stratul mai vechi de sorginte săsească. V. Arvinte[1] a arătat că termeni ca *şopron, ştreang, turn, troacă* sînt anteriori influenţei germane moderne, de după 1700. Calculul nostru sugerează că vor fi fiind şi alţii, greu de depistat din lipsa unor criterii de delimitare satisfăcătoare.

### 7. Împrumuturi sîrbocroate

Dificultatea datării cu ajutorul expresiei (12.6) a aportului lexical sîrbocroat derivă din imposibilitatea, cel puţin momentană, de a distinge în cuprinsul clasei cuvintelor cu etimon dublu bulgar şi sîrbocroat (v. lista din paragraful 10.4) între unităţile ce prezintă într-adevăr fenomenul de etimologie dublă şi cele ce descind de fapt din elemente vechi slave („bulgare vechi"). În ipoteza că toate aceste cuvinte aparţin primei categorii, avem $m = 0{,}349$ şi $t_m = 740 \div 780$, adică rezultatul *absolut neverosimil* că, în medie, termenii de origine sîrbocroată ar data din perioada $1170 \div 1210$. În cazul contrar: $m = 0{,}587$ şi $t_m = 260 \div 275$, ceea ce împinge centrul de greutate al influenţei sîrbocroate pînă în anii $1675 \div 1690$. E de presupus că adevărul se află pe undeva pe la mijloc. La vremea sa, Ovid Densuşianu aprecia că un număr însemnat de termeni sîrbocroaţi (v. lista în ILRI p. 232—236) au pătruns în româneşte pînă în secolul al XIV-lea. Cercetări ulterioare au arătat însă că o bună parte dintre aceştia sînt de provenienţă fie slavă veche, fie bulgărească[2]. Practic, influenţa sîrbocroată asupra limbii române devine sensibilă abia începînd de prin secolul al XV-lea, cînd ea se exercită atît pe cale cultă (slavona „sîrbă" ia locul celei „bulgare" în documentele cancelariei domneşti din Ţara Românească), cît şi prin contactul direct dintre populaţii, prilejuit de imigrarea unor mase compacte de sîrbi fugiţi peste Dunăre de frica turcilor (numai în 1481 s-au strămutat astfel în Banat vreo 50.000 de sîrbi[3]). Soluţia corespunzătoare pantei $m = 0{,}349$ este deci cu totul exclusă, ceea ce reconfirmă, dacă mai era nevoie, observaţia că multe dintre cuvintele pentru care dicţionarul propune o dublă etimologie bulgară şi sîrbocroată sînt în realitate de provenienţă slavă veche.

---

[1] *Criterii de determinare a împrumuturilor săsesti ale limbii române,* ALIL, XVI (1965), p. 97—103.

[2] D. Gămulescu *op. cit.* p. 40

[3] Idem p. 43



Scanned with OKEN Scanner

## 8. Aportul ucrainian

Desfăşurarea în două faze, una veche, responsabilă de pătrunderea în limbă a unor cuvinte cunoscute tuturor dacoromânilor, alta mai recentă, în cursul căreia împrumuturile nu au depăşit statutul de regionalisme, valabilă pentru influenţele maghiară ori sîrbocroată, se întîlneşte şi în cazul aportului ucrainian. Despre vechimea primelor contacte lingvistice româno-ucrainiene, ce urcă în timp pînă în veacul al XII-lea, am mai amintit (v. paragraful 10.12). Rezultatul $t_m = 339 \div 358$, furnizat de relaţia (12.6) pentru panta $m = 0{,}515$ a dreptei caracteristice, corespunde anilor 1590 ÷ 1610, mediind acceptabil între cele două etape ale afluxului de material lexical ucrainian (pînă în secolul al XVI-lea şi după aceea[1]).

## 9. Elementul turcesc

Dreapta caracteristică (10.9) are panta $m = 0{,}690$. Introdusă în (12.6) aceasta conduce la $t_m = 189 \div 200$. Cu alte cuvinte vîrsta medie a împrumuturilor corespunde perioadei 1750 ÷ 1760, ceea ce arată că invazia de termeni turceşti din a doua jumătate a secolului al XVIII-lea şi de la începutul celui următor a cunoscut asemenea proporţii încît a fost în măsură să compenseze vechimea mult superioară a relativ puţinelor cuvinte încorporate vocabularului limbii române înainte de epoca fanariotă.

## 10. Aportul rusesc

În acelaşi secol al confruntărilor militare ruso-turce pe teritoriul principatelor române, pare să se situeze şi centrul de greutate al influenţei lexicale ruseşti. Calculul cu formula (12.6) propune deceniul 1740 ÷ 1750 drept intervalul cel mai probabil de amplasare a acestui moment. Datarea pare cam timpurie, dacă se ţine seama că un import de termeni ruseşti s-a produs şi în secolul nostru, dar trebuie să avem

[1] V. Vascenco *Elementele slave răsăritene în limba română (Periodizarea împrumuturilor lexicale)* SCL X (1959) p. 394—408.



Scanned with OKEN Scanner

în vedere, pe de o parte, că, în marea lor majoritate, noile împrumuturi (ca *activist, colhoz, cursant, instructaj* etc.) nu sînt productive şi, ca atare, nu participă la construirea dreptei caracteristice şi, pe de altă parte, că un număr însemnat de cuvinte (predominant cele cu dublă etimologie rusă şi ucraineană sau rusă şi polonă) sînt cu mult mai vechi, datînd din perioada medievală. Influenţa acestora asupra valorii medii a vîrstei elementelor ruseşti poate fi estimată astfel:

Din tabelul 10.13.1., care sintetizează aportul lexical rusesc, să scădem contribuţia fondului de cuvinte cu dublă etimologie ruso-ucraineană şi ruso-polonă, pe considerentul că, în general, aceste elemente aparţin unei faze mai vechi a legăturilor lingvistice cu slavii de răsărit. Obţinem astfel bilanţul din tabelul 12.1. şi o dreaptă caracteristică de ecuaţie:

$$l_{n\ rus} = 4{,}950 - 0{,}755\ d \qquad (12.7)$$

*Tabelul 12.1.*

| NR. DERIVATE ≥ | 1 | 2 | 3 |
|---|---|---|---|
| Nr. cuv.-bază | 66 | 29 | 14 |
| ln (Nr. cuv.-bază) | 4,19 | 3,37 | 2,64 |

După cum era de aşteptat, panta dreptei a crescut sensibil: $m = 0{,}775$ faţă de 0,668. Noii ei valori îi corespunde, conform relaţiei 12.6., o vîrstă medie de 150—158 de ani, ceea ce înseană că baricentrul influenţei ruseşti s-a deplasat cu exact o jumătate de veac, atingînd fruntariile secolului următor (anii 1790—1800).

## 11. Influenţa latină savantă

Nesiguranţele de diagnostic la care ne-am referit într-un paragraf anterior (10.15) grevează, după cum e şi firesc, asupra siguranţei datării acestei influenţe. Introducînd în formula (12.6) valoarea $m = 0{,}738$ a pantei din ecuaţia (10.15), stabilită cu respectarea strictă a etimologiilor[1] indicate în DLRM, obţinem $t_m = 165 \div 174$ de ani. Perioada corespunzătoare acestei vîrste medii este deceniul

[1] Cu unica excepţie a eliminării controversatului cuvînt *popor*, a cărui situaţie a fost discutată în paragraful 10.15.



Scanned with OKEN Scanner

1776÷1785. În structura considerată, fondul de cuvinte preluate din latina literară este deci cu circa un secol mai vechi decît neologismele zise „franceze". Rolul cărturarilor de formaţie umanistă, de la Miron Costin şi Dimitrie Cantemir pînă la reprezentanţii Şcolii Ardelene, pentru care latina reprezenta principala sursă de constituire a unei terminologii menite să sincronizeze cultura noastră cu cea europeană, este limpede pus în lumină de acest rezultat. Chiar dacă valorile absolute calculate sînt susceptibile de revizuiri, decalajul temporal important dintre influenţa latină savantă şi cea „franceză" rămîne o realitate greu de contestat.

## 12. Împrumuturile italiene

Înrîurirea italiană asupra limbii române este şi ea cu certitudine anterioară celei franceze. Cuvinte ca *spiţer* ⟨ it. *speziario, piaţă* ⟨ it. *piazza, monedă* ⟨ veneţian *moneda* circulau la noi încă din epoca feudală[1]. Deşi DLRM le consideră de origine franceză, *lazaret* sau *parolă,* atestate în scris în secolele XVII—XVIII[2], provin şi ele neîndoielnic din italieneşte. Cîteva dintre unităţile productive de obîrşie italiană, cum ar fi de pildă *balsam* sau *drit*, au toate şansele de a data din aceeaşi vreme. Epoca brîncovenească, dominată cultural de personalitatea stolnicului Constantin Cantacuzino, fost student la Padova, s-a constituit într-un preludiu al influenţei mult mai intense exercitate în primele decenii ale secolului al XIX-lea, prin literaţi filo-italieni ca Gh. Asachi sau I. Heliade-Rădulescu, dar şi prin stabilirea în principate a unor muzicieni, medici, originari din Peninsulă. Oricum, nu surprinde faptul că prin aplicarea formulei (12.6) pentru situaţia aportului lexical italian ($m$ = 0,835) anii 1815—1820 apar drept centru de greutate al acestei înrîuriri.

## 13. Cuvintele cu etimologie necunoscută

Calculul vechimii medii a acestui material cu totul neomogen nu are sens decît în măsura în care poate da o idee despre raportul cantitativ dintre sursele vechi şi cele noi implicate în procesul

---

1 Gh. Ivănescu ILR., 1980, p. 497-498.

2 V. Şerban, I. Evseev *Vocabularul românesc contemporan* Ed. Facla, 1978, p. 247.



Scanned with OKEN Scanner

constituirii sale. Panta dreptei caracteristice fiind $m = 0{,}474$, rezultă $t_m = 400 - 423$ de ani, ceea ce subliniază încă o dată prezenţa modestă în acest corpus a cuvintelor de obîrşii străvechi.

Cu aceasta am încheiat o primă încercare de datare a împrumuturilor de diferite proveniențe[1]. Valoarea rezultatelor obţinute este, firește, relativă, aceasta în primul rînd din pricina ipotezelor ce au servit la tararea modelului. Dacă este în afara oricărei îndoieli că grosul cuvintelor de origine latină datează din periaoda ocupaţiei romane a Daciei, în schimb presupunerea că în 1950 neologismele de sursă „franceză" aveau vîrsta medie de 70—80 de ani poate părea destul de arbitrară. Se poate spune ceva mai precis în acest domeniu?

O cale ar fi compararea neologismelor din DLRM cu cele conţinute de dicţionarele româneşti din secolul trecut. Din păcate însă, aceste dicţionare sînt departe de a reflecta limba vie, ele sînt mai întotdeauna proiecte abstracte ale unei române dorite de lexicograf şi, cu precădere în domeniul neologic, cuprind un număr mare de cuvinte care nu au circulat probabil niciodată. A le considera o oglindă a felului cum vorbeau străbunicii noştri ar fi, fără doar şi poate, o greşeală şi, de aceea, utilizarea lor pentru reconstituirea succesiunii şi a ritmului de pătrundere în limbă a elementelor franceze nu pare recomandabilă.

Rămîne o a doua cale, sugerată de însuşi modelul nostru. Corelarea cu anumite evenimente istorice care au marcat evoluţia limbii poate servi la determinarea unor limite *ante* şi *post quem* ale influenţelor considerate.

Calculul precedent a indicat deceniul 1750—1760 drept centru de greutate al influenţei turceşti. Teoretic, nimic nu ne împiedică să plasăm momentul respectiv atît înainte cît şi după inventarul determinat. Există însă o frontieră peste care nu se poate trece. Începînd din momenutl păcii de la Adrianopole (1829), cînd ţările române încep să iasă de sub tutela Porţii, bilanţul influenţei turceşti nu poate fi decît negativ: termenii ce se perimează şi ies din zona activă a vocabularului îi depăşesc pe cei care, puţin probabil, vor mai fi fost importaţi după această dată. Avînd în vedere vechimea multora dintre

[1] Am omis numai substratul şi formaţiile româneşti din elemente de origine latină, din pricina anomaliilor apărute în aceste cazuri ca urmare a sărăciei materialului lexical de care dispunem: cuvintele comune cu albaneza rezultă drept mai vechi decît cele latine, iar alcătuirile româneşti — absurd de recente (am arătat în paragraful 10.19. cît credit se poate acorda parametrilor calculaţi).



Scanned with OKEN Scanner

cuvintele de origine turcească, este absolut imposibil ca centrul de greutate al acestei influenţe să depăşească anul 1800.

Or, dacă elementul turcesc, considerat în ipoteza extremă de mai sus, se substituie celui „francez" la calculul constantelor $k$ şi $\gamma$, relaţia (12.6) capătă forma:

$$t_m = \frac{65}{m^{2,261}} \tag{12.8}$$

Pentru valoarea $m = 1{,}092$ a pantei din ecuaţia dreptei caracteristice a influenţei „franceze", obţinem imediat $t_m = 53$ ani. Înseamnă că centrul de greutate al acestei înrîuriri nu poate depăşi anul 1897.

Spre a fixa influenţei respective şi un termen *post quem* să ne referim acum la situaţia elementului vechi slav. Odată ce un aport lexical slav anterior anului 800 este foarte improbabil, nu pare de crezut, avînd în vedere durata multiseculară de exercitare a acestei influenţe, ca localizarea centrului ei de greutate să fie anterioară anului 900. Folosind pentru tararea formulei acest din urmă termen, se ajunge la expresia:

$$t_m = \frac{127}{m^{1,767}} \tag{12.9}$$

Acelaşi $m = 1{,}092$ corespunzător aportului „francez" furnizează acum $t_m = 109$ ani, ceea ce indică drept termen *post quem* al centrului de greutate anul 1841. Între limitele extreme astfel determinate (1841—1893) intervalul ales anterior pentru calculul constantelor din relaţia (12.6), şi anume 1870—1880, se încadrează, zicem noi, destul de bine. El se situează, ce-i drept, în a doua jumătate a perioadei, dar trebuie observat că frontiera inferioară (anul 1841) nu prezintă acelaşi grad de credibilitate ca cea superioară (1893). Ni se pare puţin probabil ca circa 50% dintre cuvintele „franceze" pătrunse în româneşte pînă în anul 1950 să fie anterioare anului 1841, ba nici chiar anului 1850. Subdimensionarea termenului *post quem* al influenţei „franceze" constituie, mai mult ca sigur, o consecinţă a împingerii exagerate în trecut a influenţei slave, exercitate cu preponderenţă după anul 900. În forma (12.6), relaţia dintre productivitate şi vechime rămîne, prin urmare, cea mai demnă de încredere, iar rezultatele obţinute cu ajutorul ei şi sumar comentate în acest capitol pot constitui o bază de discuţie şi un punct de plecare pentru cercetări şi aprofundări viitoare.



Scanned with OKEN Scanner

## 13. Confirmarea ipotezei liniarităţii în domeniul productivităţii lexicale. Modelul $M_P$.

Am zăbovit asupra problemelor de cronologie a împrumuturilor din pricina naturii mai neobişnuite a metodei de datare propuse, deşi trebuie s-o recunoaştem, tratarea acestor aspecte nu are în economia lucrării de faţă decît caracterul unei simple digresiuni. Obiectivul principal pe care continuăm să îl urmărim este confirmarea „ipotezei liniarităţii", conjectură de sorginte empirică, sugerată de distribuţia pe origini a cuvintelor din propunerea de vocabular fundamental a lui Al. Graur (FPL). Analiza productivităţii termenilor din DLRM ne îngăduie să facem un important pas înainte pe această linie.

Reamintim că, în esenţă, „ipoteza liniarităţii" afirmă următoarele: dacă înşiruim întreg vocabularul limbii române în ordinea descrescătoare a importanţei cuvintelor, considerînd acest şir de la începutul său şi pînă într-un punct arbitrar ales, mulţimea de cuvinte astfel constituită se bucură de proprietatea că, în cuprinsul ei, numărul $n_i$ al elementelor de origine $i$ şi numărul $n_j$ al elementelor de origine $j$ satisfac o relaţie de forma:

$$ln\ n_i = a\ ln\ n_j + b \qquad (13.1)$$

oricare ar fi originile $i$ şi $j$.

În această formulare generală este vorba, precum se vede, de *importanţa* cuvintelor. Or, s-a arătat că acest concept aparent vag (în orice caz prea imprecis pentru a figura într-o relaţie matematică) depinde de cîţiva parametri cuantificabili, ca productivitatea lexicală, numărul de sensuri distincte, numărul de expresii la care participă, frecvenţa de folosire. Ne-am ocupat deocamdată de primul dintre ei. Dispunem acum nu numai de şirul ordonat după productivitate al tuturor cuvintelor din DLRM, dar şi de şiruri parţiale corespunzătoare diferitelor componente etimologice ale limbii române. Cu ajutorul acestora din urmă am stabilit o relaţie importantă, care ne permite să calculăm numărul total $n_i$ de cuvinte de origine $i$ care au generat cel puţin $d$ derivate:

$$ln\ n_i = k_i - m_i d \qquad (13.2)$$



Scanned with OKEN Scanner

unde $k_i$ şi $m_i$ sînt două constante ce caracterizează influenţa limbii $i$ asupra limbii române. Pentru un alt aport lexical de origine $j$, ecuaţia (13.2) ia forma:

$$ln\ n_j = k_j - m_j d \tag{13.3}$$

În ipoteza ordonării descrescătoare a vocabularului după productivitatea lexicală, numărul de cuvinte de origine $i$ şi, respectiv, $j$ conţinute în acea parte a şirului ai cărei termeni au dat naştere la cel puţin $d$ derivate este dat de relaţiile (13.2) şi (13.3). Legătura dintre aceste numere se obţine eliminînd argumentul $d$ din cele două ecuaţii, respectiv extrăgîndu-l dintr-o relaţie şi substituind în cealaltă expresia astfel obţinută. Rezultă:

$$ln\ n_i = \frac{m_i}{m_j} ln\ n_j + \left(k_i - \frac{m_i}{m_j} k_j\right) \tag{13.4}$$

Se observă că, dacă notăm $a = \frac{m_i}{m_j}$ şi $b = k_i - \frac{m_i}{m_j} k_j$, ecuaţia (13.4) devine identică cu (13.1) „Ipoteza liniarităţii" este astfel confirmată.

Să aplicăm acest rezultat în cazul raporturilor dintre elementul vechi slav şi fondul latin moştenit. Fie deci $i$ = slav, $j$ = latin. Din relaţiile (10.1) şi (10.3) avem $m_{sl} = 0{,}312$, $k_{sl} = 6{,}878$, $m_{lat} = 0{,}230$, $k_{lat} = 7{,}475$. Introducînd aceste valori în ecuaţia (13.4), obţinem:

$$ln\ n_{sl} = 1{,}357\ ln\ n_{lat} - 3{,}262 \tag{13,5}$$

Reprezentată grafic sub forma unei drepte (fig. 13.1), ecuaţia de mai sus caracterizează complet raporturile dintre componenta latină şi cea veche slavă a vocabularului limbii române, considerate din punctul de vedere al productivităţii lexicale. Ea diferă de dreapta similară, construită pe baza generalizării situaţiei din FPL şi din DLRM (după statistica lui Macrea, ale cărei erori le-am sesizat anterior) şi a cărei ecuaţie era[1]:

$$ln\ n_{sl} = 2{,}01\ ln\ n_{lat} - 8{,}11 \tag{13,6}$$

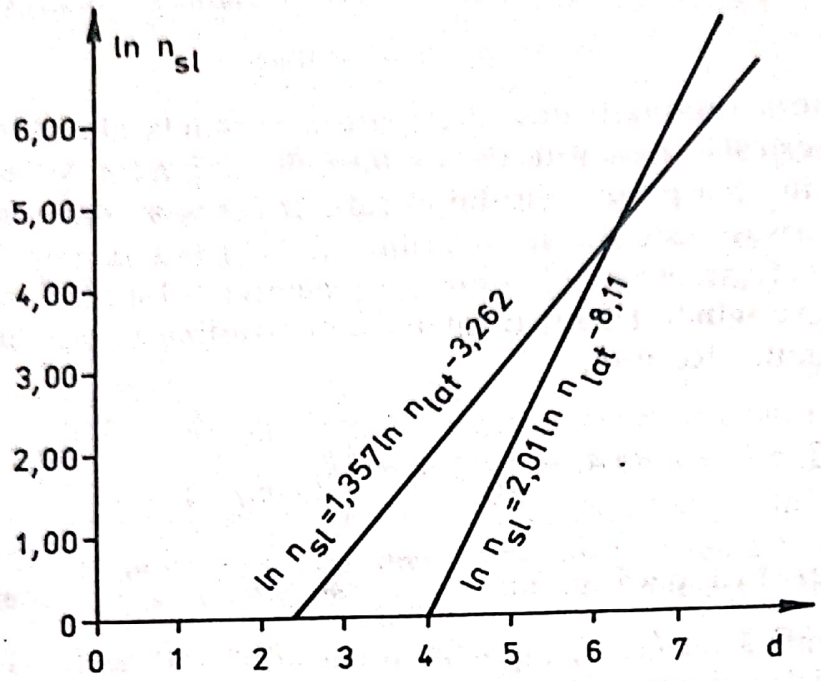
Necoincidenţa acestor două drepte (fig. 13.1) are drept cauză principală faptul că, spre deosebire de noi, Al. Graur, atunci cînd îşi selecta cuvintele pentru cele trei „cercuri" ale fondului principal, întreprindea ceea ce se numeşte o clasificare multicriterială. El ţinea

[1] V. capitolul 6, ecuaţia (6.1)



Scanned with OKEN Scanner

FIG. 13.1




implicit seama nu numai de productivitatea avută în vedere de noi în cele de mai sus, ci şi de polisemia, de gradul de utilizare şi de prezenţa în expresii şi locuţiuni ale cuvintelor. Calculul cîtorva perechi ($n_{lat}$, $n_{sl}$) de valori omoloage din relaţiile (13.5) şi (13.6) ne arată de ce fel sînt deosebirile dintre cele două clasamente.

Pentru o serie de valori ale argumentului $n_{lat}$, cuprinse între 10 şi 2000, tabelul (13.1) prezintă numerele $n_{sl}$ ce le corespund potrivit ecuaţiilor (13.5) (linia a doua a tabelului) şi (13.6) (linia a treia).

*Tabelul 13.1*

| $n_{lat}$ | 10 | 20 | 50 | 100 | 200 | 400 | 600 | 800 | 1000 | 1200 | 1400 | 1600 | 1800 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_{sl}$ (13.5) | 1 | 2 | 8 | 20 | 51 | 130 | 226 | 333 | 451 | 578 | 712 | 854 | 1002 | 1156 |
| $n_{sl}$ (13.6) | — | — | — | 3 | 13 | 51 | 115 | 206 | 322 | 465 | 633 | 832 | 1059 | 1296 |

Cum să interpretăm aceste cifre? Dacă vom parcurge lista celor mai productive cuvinte româneşti şi ne vom opri de îndată ce am ajuns la al 100-lea element de origine latină, tabelul 13.1 ne asigură că pînă în acel moment vom fi întîlnit cca. 20 de cuvinte de provenienţă slavă veche. În schimb, într-un clasament multicriterial al cuvintelor de tipul celui subiacent fondului principal lexical delimitat



Scanned with OKEN Scanner

de Al. Graur numărul de cuvinte slave întîlnite pînă al 100-lea cuvînt latin va fi de numai trei. Discrepanţa dintre cele două rezultate poate surprinde, dar numai la prima vedere. Este suficient să observăm că printre cele mai însemnate cuvinte de origine latină figurează un număr de termeni neproductivi, pronume, prepoziţii, conjuncţii, pentru ca lucrurile să se lămurească imediat. În toate aceste categorii morfologice elementul slav este inexistent. Or, atît din punctul de vedere al frecvenţei de folosire cît şi din cel al bogăţiei de funcţii sintactice[1] asemenea unităţi figurează indiscutabil printre cele mai importante cuvinte ale limbii române. Faptul că aceste elemente esenţiale ale vocabularului sînt aproape în totalitate de origine latină conferă fondului lexical moştenit o superioritate zdrobitoare asupra celorlalte componente istorice ale limbii în general şi asupra aportului slav în particular. Cu privire la acesta din urmă, este cazul să citäm statistica lui Alphonse Juilland potrivit căreia[2] printre cele mai frecvent utilizate 106 cuvinte româneşti se numără 100 de termeni proveniţi din latineşte şi numai unul de obîrşie slavă: verbul *trebui* (rangul 52). Se vede cît de mult variază imaginea pe care ne-o oferă asupra ponderii diferitelor aporturi lexicale adoptarea unor criterii diferite de apreciere. Din punctul de vedere al productivităţii, celor mai importante 100 de cuvitne latine le corespund 20 de cuvinte slave, în timp ce din punctul de vedere al frecvenţei — numai unul singur. După cum era de aşteptat, într-o clasificare ce ţine seamă de ambele criterii (precum cea „abstrasă" de noi din rezultatele lui Al. Graur), situaţia este intermediară: la 100 de elemente latineşti întîlnim 3 slave.

Extinzînd lista de cuvinte, efectul perturbator produs de submulţimea, numericeşte limitată, a termenilor importanţi, dar neproductivi se face tot mai puţin simţit şi, în consecinţă, valorile calculate cu formulele (13.5) şi (13.6) se apropie apreciabil una de alta. Există chiar un punct unde ele devin egale ($n_{lat}$ = 1676, $n_{sl}$ = 910), ceea ce înseamnă că, indiferent dacă judecăm însemnătatea cuvintelor numai după productivitatea lor sau luăm în considerare şi celelalte criterii cu care a operat Al. Graur, la 1676 de cuvinte de origine latină corespund tot 910 împrumuturi slave.

Evident, corelaţii liniare cu fondul latin precum cea stabilită pentru elementul vechi slav (13.5) pot fi construite pentru toate celelalte

---

[1] În cazul conectorilor, tipurile de conexiuni (funcţii sintactice) reprezintă analogul sensurilor cuvintelor „pline".

[2] Alphonse Juilland, P.M.H. Edwards, Ileana Juilland *Frequency Dictionary of Romanian Words* Mouton & Comp. 1965, London, The Hague, Paris.

Scanned with OKEN Scanner

componente ale vocabularului. Ansamblul lor constituie un model complet al structurii pe origini a lexicului românesc văzut prin prisma capacităţii cuvintelor de a genera termeni noi prin prefixare, sufixare sau derivare regresivă. Vom numi $M_p$ acest model axat pe productivitatea lexicală. Ecuaţiile ce îl compun sînt următoarele:

$$n_a = exp\ (0{,}852\ ln\ n_{lat} - 2{,}041)$$

$$n_{sl} = exp\ (1{,}357\ ln\ n_{lat} - 3{,}262)$$

$$n_{bg.scr} = exp\ (1{,}430\ ln\ n_{lat} - 5{,}641)$$

$$n_{onom} = exp\ (1{,}574\ ln\ n_{lat} - 6{,}092)$$

$$n_{bg} = exp\ (1{,}670\ ln\ n_{lat} - 6{,}883)$$

$$n_{magh} = exp\ (1{,}648\ ln\ n_{lat} - 6{,}860)$$

$$n_{tc} = exp\ (3.000\ ln\ n_{lat} - 16{,}191)$$

$$n_{gr} = exp\ (2{,}178\ ln\ n_{lat} - 10{,}605)$$

$$n_{germ} = exp\ (2{,}535\ ln\ n_{lat} - 13{,}644)$$

$$n_{scr} = exp\ (2{,}552\ ln\ n_{lat} - 14{,}386)$$

$$n_{lat.sav} = exp\ (3{,}265\ ln\ n_{lat} - 18{,}503)$$

$$n_{rus} = exp\ (3{,}370\ ln\ n_{lat} - 20{,}237)$$

$$n_{ucr} = exp\ (2.839\ ln\ n_{lat} - 16{,}177)$$

$$n_{it} = exp\ (3{,}630\ ln\ n_{lat} - 21{,}874)$$

$$n_{fr} = exp\ (4{,}748\ ln\ n_{lat} - 26{,}786)$$

$$n_{nec} = exp\ (2{,}061\ ln\ n_{lat} - 9{,}035)$$

$$n_{inc} = exp\ (1{,}509\ ln\ n_{lat} - 5{,}492)$$

$$n_{rom(lat)} = exp\ (3{,}109\ ln\ n_{lat} - 18{,}550)$$

$$n_{n.s.m.} = exp\ (1{,}539\ ln\ n_{lat} - 7{,}485)$$

Ele se referă, în ordine, la aporturile lexicale: autohton, vechi slav, bulgaro-sîrbocroat, onomatopeic, bulgar, maghiar, turcesc, grecesc, german, sîrbocroat, latin savant, rusesc, ucrainian, italian, francez, necunoscut, incert, la cuvintele formate în româneşte din elemente



Scanned with OKEN Scanner

latine, iar ultima ecuaţie priveşte un material lexical eterogen, provenit din limbile slave de nord (polonă, ucraineană, rusă), care grupează la un loc cuvintele cu etimologie dublă rusă şi ucraineană enumerate în paragraful 10.12, pe cele polono-ucrainiene *(coreţ, cufăr, darabană, ghint, mutelcă, poleac, povidlă, puhă, romaniţă, şleau, ţidulă, zolniţă)*, ruso-poloneze *(arendă, bunt, hoardă, nurcă, polspanc, roată* = „companie de ostaşi"), poloneze *(başcă, baştă, ceapcă, chilnă, clapon, comenduire, comornic, daraban, dardă, drăgan, dulău, harapnic, hască, hatman, herb, horbotă* etc.), polono-ruso-ucrainiene *(muscal, pihotă)*, ca şi toate celelalte unităţi cu origine dublă sau multiplă care provin şi din cel puţin una dintre aceste trei limbi, cum ar fi, de exemplu, *contăş* (⟨ bg. & pol.), *erhă* (⟨ magh. & pol.), *rabin* (⟨ germ. & pol), *panţir* ( ⟨ pol. & scr.), *plită* (⟨ ceh. & rus. & ucr.), *garniţă* (⟨ pol. & rus & scr.), *proviant* (⟨ germ. & pol. & rus), *ofiţer* (⟨ fr. & pol. & rus), *şafăr* (⟨ germ. & pol. & ucr.), *posteucă* (⟨ bg. & rus. & ucr.), *podină* (⟨ rus & scr. & ucr.), *poleac* (⟨ pol. & scr. & ucr.), *palancă* (⟨ magh. & pol. & tc.), *puşcă* (⟨ magh. & pol. & rus), *ba* (⟨ bg. & scr. & pol. & ucr.), *podnojă* (⟨ bg. & scr. & rus & ucr.), *şcoală* (⟨ bg. & scr. & pol. & rus). Rostul acestei categorii pestriţe este de a acoperi cît mai fără rest ansamblul cuvintelor din dicţionar. Fiecare dintre subclasele incluse aici se dovedea numericeşte neîndestulătoare pentru a ne permite construirea unei ecuaţii ad-hoc. Omiterea lor ar fi mărit însă decalajul dintre modelul matematic şi compoziţia reală a lexicului românesc. Pe de altă parte, cuvintele cu etimologie dublă sau multiplă nu pot fi incluse în clasele indicate de dicţionar, căci în acest caz ar figura în model de două sau mai multe ori. De aceea, de pildă, ecuaţiile corespunzătoare elementelor rus şi ucrainian fac abstracţie de existenţa cuvintelor cu etimon dublu ucrainian şi rus, acestea din urmă fiind înglobate în categoria împrumuturilor pe care le-am numit *nord-slave medievale* şi la care se referă ultima relaţie a modelului. Denumirea de *nord-slave* se justifică prin existenţa pentru fiecare dintre cuvintele conţinute în această mulţime a cel puţin unui etimon fie polonez, fie rus, fie ucrainian, iar epitetul *medieval* trebuie pus în legătură cu vechimea medie apreciabilă pe care o atestă forma ecuaţiei caracteristice corespunzătoare acestui corpus:

$$ln\, n_{n.s.m.} = 4{,}020 - 0{,}354\, d$$

Într-adevăr, pentru panta $m_{n.s.m.} = 0{,}354$, formula (12.6) ne indică o vîrstă mijlocie de 720—760 de ani. Evident, acest rezultat nu posedă credibilitatea altora similare, întrucît relaţia (12.6) este valabilă numai

Scanned with OKEN Scanner

pentru mulţimi de cuvinte etimologic omogene, ceea ce nu e deloc cazul aici. Observaţia că elementele acestui corpus eteroclit sînt în general vechi rămîne însă în picioare, chiar dacă vârsta calculată e probabil exagerat de mare.

## 14. Explicitarea modelului Mp. Raporturile numerice dintre cuvintele de diferite proveniențe

Principalul cîştig obţinut prin elaborarea modelului Mp este că, mulţumită lui, putem pentru prima oară depăşi viziunea unilaterală, trunchiată, asupra rolului diferitelor surse etimologice în constituirea vocabularului limbii române pe care ne-o ofereau statisticile bazate pe un corpus lexical determinat (fond principal, vocabular fundamental, vocabular reprezentativ[1] sau chiar totalitatea cuvintelor dintr-un dicţionar, ca în cercetările lui D. Macrea). Neconcordanţele dintre statisticile efectuate de cercetători diferiţi îşi aveau sursa, după cum am mai avut prilejul să o arătăm (v. paragraful 7), nu atît în modul cum fusese aleasă mulţimea de cuvinte supusă analizei, cît în faptul că o asemenea mulţime a fost la urma urmei aleasă, deoarece ştim acum că nici un corpus lexical particular nu poate reflecta structura de ansamblu a vocabularului limbii române.

Cele mai importante 100 de cuvinte româneşti prezintă o cu totul altă distribuţie pe origini decît cele mai importante 1 000 de cuvinte, care au, la rândul lor, o repartiţie etimologică mult diferită de a celor mai însemnate 10 000 de cuvinte şi aşa mai departe. Am mai evidenţiat acest aparent paradox; să vedem acum în ce fel poate fi el depăşit.

Soluţia este pe cît de simplă pe atît de greu de pus în practică. O imagine cît mai corectă a situaţiei întregului vocabular presupune efectuarea a cît mai multe secţiuni statistice prin diversele straturi ale piramidei lexicale. Firește, bilanţurile întocmite de Al. Graur pentru cele 965, 1181 sau 1419 cuvinte din „fondul principal lexical", de

[1] Cercetarea mai nouă a unui colectiv coordonat de Marius Sala (*Vocabularul reprezentativ al limbilor romanice* Ed. Ştiinţifică şi Enciclopedică, Bucureşti, 1988) suferă şi ea de acelaşi cusur.



Scanned with OKEN Scanner

colectivul condus de M. Sala pentru cele 2 558 de cuvinte din „vocabularul reprezentativ", de D. Macrea pentru dicţionarele explicative din 1931 („Cartea Românească") şi 1958 (DLRM) sînt corecte şi utile, dar ele ne prezintă fiecare în parte cîte o singură faţetă a realităţii proteice care este lexicul românesc. Doar prin multiplicarea la nesfîrşit a unor atari incursiuni (numai autorii lor ştiu cît de laborioase şi delicate) s-ar putea ajunge cîndva, din aproape în aproape, la o caracterizare completă a raportului de forţe dintre diferitele componente etimologice ale vocabularului limbii române.

Indiscutabil, o asemenea cale nu este de recomandat. Rămîne o a doua soluţie, cea a modelării matematice, adică a descoperirii unui sistem de relaţii analitice care să descrie cît mai fidel structura vocabularului. Odată elaborat acest model, efectuarea unui număr oricît de mare de secţiuni de tipul celor realizate cu atîta trudă de către autorii citaţi devine o problemă algoritmizabilă ce se rezolvă fără dificultăţi cu ajutorul unui calculator electronic. Am încercat să facem acest lucru pentru modelul $M_p$, rezultatele fiind prezentate (selectiv, din economie de spaţiu tipografic) în tabelul 14.1.

Fiecare linie a tabelului reprezintă o secţiune distinctă prin edificiul vocabularului, ce oferă răspuns la întrebarea: care este distribuţia pe origini a celor mai productive $n$ cuvinte româneşti? Pentru a acoperi practic toate cazurile interesante, domeniul de variaţie al numărului $n$ s-a ales între zero şi 45 000, limita superioară depăşind de cca. 7 ori numărul de cuvinte productive din DLRM. Avînd în vedere cerinţa firească de detaliere a straturilor mai adînci ale vocabularului, sediu al cuvintelor mai importante, mai stabile, treptele de variaţie ale lui $n$ s-au ales diferenţiat după cum urmează:

— între $n = 0$ şi $n = 100$ din 10 în 10
— între $n = 100$ şi $n = 1\,000$ din 20 în 20
— între $n = 1\,000$ şi $n = 2\,000$ din 50 în 50
— între $n = 2\,000$ şi $n = 5\,000$ din 200 în 200
— între $n = 5\,000$ şi $n = 10\,000$ din 1 000 în 1 000
— între $n = 10\,000$ şi $n = 45\,000$ din 5 000 în 5 000

În plus, au fost introduse linii speciale de cîte ori îşi făcea apariţia un element nou, recte un cuvînt de o origine încă neîntîlnită pînă în acel moment. Pentru mai uşoara identificare a acestor situaţii, numărul $n$ corespunzător figurează îngroşat în prima coloană. În acelaş scop, ordinea originilor indicate de coloane este cea a ocurenţei primului element de provenienţa respectivă. Comentariile ce succed tabelului vor face mai lesne de înţeles rostul tuturor acestor convenţii.



Scanned with OKEN Scanner

*Tabelul 14.1..*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | INC. | ONOM. | BG. & SCR | BG. | MAGH. | NEC. | N.S.M. | GR. | GERM. | TC. | SCR. | LAT. SAV. | UCR. | „FR.." | ROM (LAT) | RUS | ITAL. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 4 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 9 | 7 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 10 | 8 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 20 | 17 | 1 | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 30 | 25 | 2 | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 31 | 25 | 2 | 3 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 39 | 31 | 2 | 4 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 40 | 31 | 2 | 4 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 50 | 38 | 3 | 6 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 54 | 41 | 3 | 6 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 57 | 43 | 3 | 6 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 60 | 45 | 3 | 7 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 70 | 52 | 4 | 8 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 78 | 58 | 4 | 9 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 80 | 59 | 4 | 10 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 90 | 66 | 5 | 11 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 100 | 72 | 5 | 13 | 3 | 2 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 117 | 83 | 6 | 15 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 120 | 85 | 6 | 16 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 135 | 95 | 6 | 18 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 140 | 98 | 6 | 19 | 4 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 160 | 110 | 7 | 22 | 5 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 180 | 123 | 8 | 26 | 6 | 4 | 3 | 3 | 3 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 200 | 135 | 8 | 29 | 7 | 5 | 4 | 4 | 3 | 3 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 220 | 147 | 9 | 33 | 8 | 6 | 4 | 4 | 4 | 3 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 240 | 158 | 10 | 37 | 8 | 6 | 5 | 5 | 4 | 4 | 1 | 2 | — | — | — | — | — | — | — | — | — |
| 254 | 166 | 10 | 39 | 9 | 7 | 5 | 5 | 5 | 4 | 1 | 2 | 1 | — | — | — | — | — | — | — | — |
| 260 | 169 | 10 | 40 | 9 | 7 | 5 | 5 | 5 | 5 | 2 | 2 | 1 | — | — | — | — | — | — | — | — |
| 276 | 176 | 11 | 43 | 10 | 8 | 6 | 6 | 5 | 5 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — |
| 280 | 179 | 11 | 44 | 10 | 8 | 6 | 6 | 5 | 5 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — |
| 300 | 191 | 11 | 47 | 11 | 9 | 6 | 7 | 6 | 6 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — |
| 320 | 201 | 12 | 51 | 12 | 9 | 7 | 7 | 7 | 7 | 2 | 3 | 1 | 1 | — | — | — | — | — | — | — |
| 340 | 213 | 12 | 55 | 13 | 10 | 8 | 8 | 7 | 7 | 2 | 3 | 1 | 1 | — | — | — | — | — | — | — |
| 347 | 214 | 13 | 56 | 14 | 11 | 8 | 8 | 7 | 8 | 2 | 3 | 1 | 1 | 1 | — | — | — | — | — | — |
| 353 | 218 | 13 | 57 | 14 | 11 | 8 | 8 | 7 | 8 | 2 | 3 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 360 | 222 | 13 | 59 | 14 | 11 | 8 | 8 | 8 | 8 | 2 | 3 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 380 | 233 | 13 | 62 | 15 | 12 | 9 | 9 | 8 | 9 | 2 | 4 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 384 | 234 | 14 | 63 | 15 | 12 | 9 | 9 | 8 | 9 | 2 | 4 | 1 | 1 | 1 | 1 | 1 | — | — | — | — |
| 400 | 242 | 14 | 65 | 16 | 13 | 9 | 10 | 9 | 10 | 3 | 4 | 1 | 1 | 1 | 1 | 1 | — | — | — | — |
| 405 | 244 | 14 | 67 | 16 | 13 | 9 | 10 | 9 | 10 | 3 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | — | — | — |
| 420 | 252 | 14 | 69 | 17 | 14 | 10 | 10 | 10 | 11 | 3 | 4 | 1 | 1 | 1 | 1 | 1 | 1 | — | — | — |



Scanned with OKEN Scanner

*Tabelul 14.1 (continuare)*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | INC. | ONOM. | BG. & SCR | BG. | MAGH. | NEC. | N.S.M. | GR. | GERM. | TC. | SCR. | LAT. SAV. | UCR. | „FR." | ROM. (LAT.) | RUS. | ITAL. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 440 | 262 | 15 | 73 | 18 | 14 | 10 | 11 | 10 | 11 | 3 | 5 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| 460 | 272 | 15 | 77 | 19 | 15 | 11 | 12 | 11 | 12 | 3 | 5 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| 480 | 282 | 16 | 81 | 21 | 16 | 11 | 13 | 11 | 13 | 3 | 5 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| 500 | 292 | 16 | 85 | 22 | 17 | 12 | 13 | 12 | 14 | 3 | 6 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| 520 | 301 | 17 | 88 | 23 | 18 | 12 | 14 | 13 | 15 | 4 | 6 | 2 | 3 | 1 | 1 | 1 | 1 | — | — | — |
| 540 | 309 | 17 | 92 | 24 | 19 | 13 | 15 | 13 | 16 | 4 | 7 | 2 | 3 | 1 | 2 | 1 | 2 | — | — | — |
| **546** | 312 | 17 | 93 | 24 | 19 | 13 | 15 | 14 | 16 | 4 | 7 | 2 | 3 | 1 | 2 | 1 | 2 | 1 | — | — |
| 560 | 319 | 18 | 95 | 25 | 20 | 13 | 15 | 14 | 17 | 4 | 7 | 3 | 3 | 1 | 2 | 1 | 2 | 1 | — | — |
| 580 | 329 | 18 | 99 | 26 | 21 | 14 | 16 | 15 | 18 | 4 | 7 | 3 | 3 | 1 | 2 | 1 | 2 | 1 | — | — |
| **585** | 330 | 18 | 100 | 26 | 21 | 14 | 16 | 15 | 18 | 4 | 8 | 3 | 3 | 2 | 2 | 1 | 2 | 1 | 1 | — |
| 600 | 336 | 18 | 103 | 27 | 22 | 15 | 17 | 15 | 19 | 4 | 8 | 3 | 4 | 2 | 2 | 1 | 2 | 1 | 1 | — |
| **613** | 342 | 19 | 105 | 27 | 22 | 15 | 17 | 16 | 20 | 4 | 8 | 3 | 4 | 2 | 2 | 1 | 3 | 1 | 1 | 1 |
| 620 | 344 | 19 | 106 | 28 | 22 | 15 | 18 | 16 | 20 | 5 | 8 | 3 | 4 | 2 | 2 | 2 | 3 | 1 | 1 | 1 |
| 640 | 353 | 19 | 110 | 29 | 23 | 16 | 19 | 17 | 21 | 5 | 9 | 3 | 4 | 2 | 2 | 2 | 3 | 1 | 1 | 1 |
| 660 | 361 | 20 | 114 | 30 | 24 | 16 | 19 | 17 | 22 | 5 | 9 | 4 | 4 | 2 | 3 | 2 | 3 | 1 | 1 | 1 |
| 680 | 370 | 20 | 117 | 31 | 25 | 17 | 20 | 18 | 24 | 5 | 10 | 4 | 5 | 2 | 3 | 2 | 4 | 1 | 1 | 1 |
| 700 | 381 | 20 | 122 | 32 | 26 | 17 | 21 | 19 | 25 | 5 | 10 | 4 | 5 | 2 | 3 | 2 | 4 | 1 | 1 | 1 |
| 720 | 389 | 21 | 125 | 33 | 27 | 18 | 22 | 19 | 26 | 5 | 11 | 4 | 5 | 2 | 3 | 2 | 5 | 1 | 1 | 1 |
| 740 | 398 | 21 | 129 | 34 | 28 | 18 | 22 | 20 | 27 | 6 | 11 | 5 | 6 | 2 | 3 | 2 | 5 | 1 | 1 | 1 |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 760 | 404 | 22 | 132 | 35 | 29 | 19 | 23 | 21 | .28 | 6 | 12 | 5 | 6 | 3 | 4 | 2 | 6 | 1 | 1 | 1 |
| 780 | 412 | 22 | 136 | 37 | 30 | 20 | 24 | 21 | 29 | 6 | 12 | 5 | 7 | 3 | 4 | 3 | 6 | 1 | 1 | 1 |
| 800 | 421 | 22 | 139 | 38 | 31 | 20 | 25 | 22 | 31 | 6 | 13 | 5 | 7 | 3 | 4 | 3 | 7 | 1 | 1 | 1 |
| 820 | 430 | 23 | 143 | 39 | 31 | 21 | 26 | 23 | 32 | 6 | 13 | 6 | 7 | 3 | 4 | 3 | 7 | 1 | 1 | 1 |
| 840 | 436 | 23 | 147 | 40 | 32 | 21 | 26 | 24 | 33 | 7 | 14 | 6 | 8 | 3 | 5 | 3 | 8 | 1 | 1 | 1 |
| 860 | 446 | 23 | 150 | 41 | 33 | 22 | 27 | 24 | 34 | 7 | 15 | 6 | 8 | 3 | 5 | 3 | 9 | 2 | 1 | 1 |
| 880 | 453 | 24 | 154 | 42 | 34 | 22 | 28 | 25 | 36 | 7 | 15 | 6 | 9 | 3 | 5 | 3 | 10 | 2 | 1 | 1 |
| 900 | 460 | 24 | 157 | 43 | 35 | 23 | 29 | 26 | 37 | 7 | 16 | 7 | 9 | 4 | 5 | 3 | 10 | 2 | 2 | 1 |
| 920 | 467 | 24 | 161 | 44 | 36 | 23 | 30 | 26 | 38 | 7 | 16 | 7 | 10 | 4 | 6 | 4 | 11 | 2 | 2 | 2 |
| 940 | 475 | 25 | 165 | 45 | 37 | 24 | 30 | 27 | 39 | 7 | 17 | 7 | 10 | 4 | 6 | 4 | 12 | 2 | 2 | 2 |
| 960 | 483 | 25 | 168 | 46 | 38 | 24 | 31 | 28 | 41 | 8 | 17 | 8 | 10 | 4 | 6 | 4 | 13 | 2 | 2 | 2 |
| 980 | 491 | 25 | 171 | 47 | 39 | 25 | 32 | 28 | 42 | 8 | 18 | 8 | 11 | 4 | 7 | 4 | 14 | 2 | 2 | 2 |
| 1000 | 498 | 26 | 175 | 48 | 40 | 26 | 33 | 29 | 43 | 8 | 19 | 8 | 11 | 4 | 7 | 4 | 15 | 2 | 2 | 2 |
| 1050 | 515 | 27 | 184 | 51 | 42 | 27 | 35 | 31 | 46 | 8 | 20 | 9 | 13 | 5 | 8 | 5 | 18 | 2 | 2 | 2 |
| 1100 | 533 | 27 | 193 | 54 | 44 | 28 | 37 | 33 | 50 | 9 | 22 | 10 | 14 | 5 | 9 | 5 | 21 | 2 | 2 | 2 |
| 1150 | 550 | 28 | 201 | 56 | 47 | 30 | 39 | 35 | 53 | 9 | 23 | 11 | 16 | 6 | 10 | 6 | 24 | 2 | 2 | 2 |
| 1200 | 567 | 29 | 209 | 59 | 49 | 31 | 41 | 36 | 56 | 10 | 25 | 11 | 17 | 6 | 11 | 6 | 28 | 3 | 3 | 3 |
| 1250 | 583 | 30 | 217 | 62 | 51 | 32 | 43 | 38 | 60 | 10 | 26 | 12 | 19 | 6 | 12 | 7 | 32 | 4 | 3 | 3 |
| 1300 | 599 | 30 | 226 | 64 | 53 | 33 | 45 | 40 | 63 | 11 | 28 | 13 | 20 | 7 | 13 | 7 | 36 | 4 | 4 | 4 |
| 1350 | 616 | 31 | 233 | 67 | 55 | 35 | 47 | 41 | 67 | 11 | 29 | 14 | 22 | 7 | 14 | 8 | 41 | 4 | 4 | 4 |
| 1400 | 632 | 32 | 242 | 69 | 58 | 36 | 49 | 43 | 70 | 11 | 31 | 15 | 23 | 8 | 15 | 8 | 46 | 4 | 4 | 4 |
| 1450 | 645 | 32 | 249 | 72 | 60 | 37 | 51 | 45 | 74 | 12 | 33 | 16 | 25 | 8 | 16 | 9 | 51 | 5 | 5 | 5 |



Scanned with OKEN Scanner

*Tabelul 14.1 (continuare)*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | INC. | ONOM. | BG. & SCR | BG. | MAGH. | NEC. | N.S.M. | GR. | GERM. | TC. | SCR. | LAT. SAV. | UCR. | „FR." | ROM. (LAT.) | RUS. | ITAL. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1500 | 660 | 33 | 257 | 74 | 62 | 38 | 53 | 47 | 77 | 12 | 34 | 17 | 27 | 9 | 18 | 10 | 57 | 5 | 5 | 5 |
| 1550 | 674 | 33 | 265 | 77 | 64 | 39 | 54 | 48 | 81 | 13 | 36 | 18 | 29 | 9 | 19 | 10 | 63 | 6 | 6 | 6 |
| 1600 | 689 | 34 | 272 | 79 | 66 | 41 | 56 | 50 | 84 | 13 | 38 | 19 | 30 | 10 | 20 | 11 | 70 | 6 | 6 | 6 |
| 1650 | 703 | 35 | 280 | 81 | 68 | 42 | 58 | 52 | 88 | 14 | 39 | 20 | 32 | 10 | 21 | 11 | 77 | 6 | 6 | 7 |
| 1700 | 717 | 35 | 287 | 84 | 70 | 43 | 60 | 53 | 91 | 14 | 41 | 20 | 34 | 11 | 23 | 12 | 84 | 7 | 7 | 7 |
| 1750 | 729 | 36 | 294 | 86 | 73 | 44 | 62 | 55 | 95 | 14 | 43 | 22 | 36 | 11 | 24 | 13 | 91 | 7 | 7 | 8 |
| 1800 | 743 | 36 | 301 | 89 | 75 | 45 | 64 | 56 | 98 | 15 | 44 | 22 | 38 | 12 | 25 | 13 | 100 | 8 | 8 | 8 |
| 1850 | 755 | 37 | 308 | 91 | 77 | 46 | 66 | 58 | 102 | 15 | 46 | 23 | 40 | 12 | 27 | 14 | 108 | 8 | 8 | 9 |
| 1900 | 768 | 37 | 315 | 93 | 79 | 47 | 68 | 60 | 105 | 15 | 48 | 24 | 42 | 13 | 28 | 15 | 117 | 8 | 9 | 9 |
| 1950 | 780 | 38 | 322 | 95 | 81 | 49 | 69 | 61 | 109 | 16 | 49 | 25 | 44 | 14 | 30 | 15 | 126 | 9 | 9 | 9 |
| 2000 | 792 | 38 | 329 | 98 | 83 | 50 | 71 | 63 | 112 | 16 | 51 | 26 | 46 | 14 | 31 | 16 | 135 | 9 | 9 | 11 |
| 2200 | 837 | 40 | 354 | 106 | 90 | 54 | 78 | 69 | 126 | 18 | 58 | 31 | 55 | 16 | 38 | 19 | 176 | 11 | 11 | 13 |
| 2400 | 880 | 42 | 379 | 114 | 97 | 58 | 85 | 75 | 140 | 19 | 64 | 35 | 63 | 18 | 44 | 22 | 223 | 13 | 14 | 15 |
| 2600 | 920 | 44 | 403 | 122 | 105 | 61 | 91 | 80 | 153 | 20 | 71 | 39 | 72 | 21 | 51 | 24 | 275 | 14 | 16 | 18 |
| 2800 | 958 | 45 | 425 | 130 | 111 | 65 | 97 | 86 | 166 | 22 | 77 | 43 | 81 | 23 | 58 | 27 | 331 | 16 | 18 | 21 |
| 3000 | 992 | 46 | 446 | 137 | 118 | 68 | 103 | 91 | 179 | 23 | 83 | 47 | 91 | 25 | 65 | 30 | 394 | 18 | 20 | 24 |
| 3200 | 1025 | 48 | 467 | 144 | 124 | 72 | 109 | 96 | 191 | 24 | 89 | 51 | 100 | 27 | 72 | 33 | 458 | 20 | 23 | 27 |
| 3400 | 1056 | 49 | 486 | 151 | 131 | 75 | 115 | 101 | 204 | 25 | 95 | 55 | 109 | 29 | 79 | 36 | 527 | 22 | 25 | 30 |
| 3600 | 1086 | 50 | 504 | 157 | 136 | 78 | 120 | 105 | 215 | 26 | 101 | 59 | 119 | 32 | 86 | 39 | 602 | 24 | 28 | 33 |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3800 | 1113 | 51 | 522 | 162 | 140 | 81 | 125 | 110 | .226 | 27 | 107 | 63 | 128 | 34 | 94 | 42 | 638 | 26 | 30 | 36 |
| 4000 | 1139 | 52 | 538 | 169 | 147 | 83 | 131 | 114 | 238 | 28 | 113 | 67 | 138 | 36 | 101 | 45 | 760 | 28 | 33 | 40 |
| 4200 | 1166 | 53 | 555 | 175 | 152 | 86 | 135 | 119 | 249 | 29 | 118 | 70 | 147 | 38 | 108 | 48 | 844 | 30 | 35 | 43 |
| 4400 | 1189 | 54 | 570 | 180 | 157 | 89 | 140 | 123 | 259 | 30 | 124 | 74 | 157 | 40 | 116 | 51 | 931 | 32 | 38 | 46 |
| 4600 | 1212 | 55 | 587 | 185 | 161 | 91 | 144 | 126 | 269 | 31 | 129 | 78 | 166 | 42 | 123 | 54 | 1024 | 34 | 40 | 49 |
| 4800 | 1235 | 56 | 601 | 191 | 166 | 94 | 149 | 131 | 281 | 32 | 134 | 82 | 175 | 44 | 131 | 56 | 1110 | 36 | 43 | 53 |
| 5000 | 1257 | 57 | 615 | 195 | 170 | 96 | 153 | 134 | 291 | 33 | 139 | 85 | 185 | 46 | 138 | 59 | 1208 | 38 | 45 | 56 |
| 6000 | 1353 | 60 | 680 | 218 | 191 | 106 | 173 | 151 | 339 | 37 | 163 | 102 | 230 | 55 | 175 | 73 | 1716 | 47 | 58 | 73 |
| 7000 | 1435 | 64 | 737 | 239 | 210 | 116 | 192 | 167 | 381 | 41 | 186 | 119 | 274 | 64 | 211 | 86 | 2259 | 57 | 71 | 91 |
| 8000 | 1505 | 66 | 786 | 257 | 227 | 124 | 208 | 181 | 422 | 44 | 207 | 135 | 317 | 73 | 247 | 99 | 2841 | 67 | 84 | 110 |
| 9000 | 1569 | 69 | 831 | 274 | 242 | 132 | 223 | 194 | 459 | 47 | 225 | 150 | 359 | 81 | 228 | 112 | 3462 | 74 | 93 | 122 |
| 10000 | 1625 | 71 | 872 | 288 | 256 | 139 | 236 | 205 | 494 | 49 | 244 | 164 | 399 | 88 | 315 | 123 | 4089 | 86 | 110 | 147 |
| 15000 | 1848 | 79 | 1038 | 350 | 314 | 166 | 292 | 253 | 644 | 60 | 323 | 226 | 587 | 122 | 476 | 177 | 7528 | 125 | 165 | 227 |
| 20000 | 2009 | 85 | 1164 | 398 | 358 | 188 | 337 | 291 | 765 | 68 | 388 | 281 | 754 | 152 | 623 | 225 | 11210 | 166 | 223 | 315 |
| 25000 | 2138 | 89 | 1265 | 436 | 394 | 205 | 373 | 323 | 869 | 75 | 444 | 328 | 908 | 178 | 760 | 268 | 15083 | 200 | 273 | 391 |
| 30000 | 2245 | 93 | 1352 | 470 | 426 | 220 | 405 | 350 | 962 | 81 | 493 | 371 | 1052 | 202 | 890 | 308 | 19041 | 247 | 323 | 469 |
| 35000 | 2340 | 96 | 1430 | 500 | 454 | 233 | 434 | 374 | 1047 | 86 | 540 | 413 | 1191 | 224 | 1017 | 346 | 23099 | 264 | 370 | 542 |
| 40000 | 2420 | 99 | 1497 | 526 | 479 | 245 | 459 | 396 | 1123 | 91 | 581 | 449 | 1318 | 244 | 1132 | 381 | 27256 | 290 | 410 | 604 |
| 45000 | 2495 | 102 | 1560 | 551 | 503 | 256 | 483 | 416 | 1195 | 95 | 621 | 485 | 1444 | 264 | 1249 | 416 | 31410 | 320 | 456 | 679 |



Scanned with OKEN Scanner

Potrivit tabelului 14.1, ierarhia productivităţii acordă prioritate absolută fondului lexical latin. Cuvintele moştenite din latineşte nu numai că ocupă primele locuri ale clasamentului, dar reprezintă chiar majoritatea absolută pînă aproape de $n = 1000$, (dintre cele mai productive 980 de cuvinte, 491 sînt latine) şi pe cea relativă pînă după $n = 5000$ (cînd ajung a fi depăşite de neologismele „franceze"). Dintre primele 10 cuvinte, 8 sînt latine (80%), dintre primele 50 – 38 (76%), dintre primele 100 – 72 (72%). Odată cu coborîrea către ranguri inferioare, procentul scade treptat în favoarea aporturilor alogene. Cel care secundează direct elementul latin este substratul: între primele 5 cuvinte productive figurează, alături de 4 latine, şi unul comun cu albaneza. Curînd însă, elementul autohton cedează pasul în faţa superstratului slav. Dacă la nivelul celor mai productive 10 cuvinte, ponderea aportului vechi slav este practic identică cu cea a substratului, ea ajunge, pentru $n = 3600$, de 10 ori mai însemnată: 504 unităţi slave faţă de numai 50 comune cu albaneza.

Poziţia elementului slav mai este consolidată şi de apariţia timpurie (la rangul 40) a primului termen de obîrşie dublă bulgaro-sîrbocroată.

Pînă la nivelul primelor 100 de unităţi lexicale, mai asistăm la apariţia, în ordine, a cuvintelor de origine onomatopeică (rangul 39), bulgărească (rangul 54) şi maghiară (rangul 57). Prezenţa în lotul acestor 100 de cuvinte a trei termeni de origine nesigură şi a unuia cu etimologie necunoscută ne avertizează că tabloul obţinut nu este definitiv, el putînd fi încă uşor modificat pe măsura progreselor cercetării etimologice. Susceptibile de îmbogăţire sînt probabil clasa elementelor de substrat (prin preluarea unor unităţi cu obârşie momentan necunoscută) şi, cu şanse apropiate, cele ale unităţilor latine, vechi slave şi maghiare (v. lista de etimologii nesigure de la pag. 137—139).

Pe parcursul „centuriei" următoare ($n = 101 - 200$), se mai ivesc cîte un cuvînt „nord-slav medieval" (rangul 117) şi unul grecesc (rangul 135). Alte două elemente, cel german (rangul 254) şi cel turcesc (rangul 276) apar în intervalul $n = 201 - 300$. Ocurenţa aporturilor sîrbocroat (rangul 347), latin savant (rangul 353) şi ucrainian (rangul 384) reprezintă grupul de evenimente notabile din zona $n = 301 - 400$. După rangul 405, care marchează apariţia importantului element „francez", nu mai sînt de semnalat decît două influenţe ce n-au reuşit să-şi trimită reprezentanţi în lotul primelor 500 de cuvinte: cea rusă ($n = 585$) şi cea italiană ($n = 613$).

Succesiunea intrării în scenă a diferitelor aporturi nu este însă nici pe departe suficientă pentru caracterizarea locului ocupat de acestea în alcătuirea vocabularului, căci prima ocurenţă a unui cuvînt de o anumită origine constituie de-abia semnalul declanşator al unei aprige



Scanned with OKEN Scanner

competiţii pentru întîietatea numerică, luptă ce ia proporţii pe măsura extinderii cercului lexical considerat. Ritmul de creştere al unora dintre componente fiind mai rapid decît al altora, vom fi martorii unei întregi serii de răsturnări de situaţie. Astfel, dacă la nivelul celor mai productive 1 500 de cuvinte, împrumuturile „franceze" nu reprezentau nici o zecime din fondul latin (57 versus 660), în mulţimea de 40 000 de cuvinte defalcată pe origini în penultima linie a tabelului 14.1., ele au ajuns, exact dimpotrivă, de peste 10 ori mai multe (27 256 faţă de 2 420). Că lucrurile stau într-adevăr aşa ne-o confirmă acele statistici care au îmbrăţişat conţinutul unui întreg dicţionar. S-ar putea replica totuşi că un raport atît de defavorabil elementului latin moştenit nu a înregistrat nici D. Macrea, dar trebuie spus că aceasta s-a întîmplat numai fiindcă în cercetările sale D. Macrea nu a întîlnit chiar 40 000 de cuvinte productive, ci numai vreo 6—7 000. Este de altfel foarte probabil ca limba română să nici nu posede un fond productiv atît de întins, dar modelarea matematică are, între alte avantaje, şi pe acela de a permite extrapolări, ceea ce în cazul de faţă s-ar traduce prin posibilitatea elaborării unor scenarii de tipul: „cum s-ar fi modificat raportul de forţe dintre componentele limbii române în ipoteza extinderii vocabularului la 200 000, 500 000 sau 1 milion de cuvinte?" Nu încape îndoială că, prin lărgirea la maximum a cercului lexical considerat, corpusul respectiv s-ar îmbogăţi cu zecile de mii de termeni prezenţi pînă acum doar în dicţionarele de neologisme şi în cele de diverse specialităţi, în care aportul „francez" este covîrşitor, astfel încît previziunile cuprinse în ultimele linii ale tabelului privind evoluţia raporturilor dintre elementul latin şi cel „francez" devin perfect plauzibile.

În general, studiul dinamicii comparative a cuplurilor de influenţe lexicale merită mai mult decît o menţiune. El se înscrie într-o problematică deja enunţată (v. capitolul 7) odată cu punerea în evidenţă a ceea ce am numit „paradoxul lui A. de Cihac". Am arătat cu acel prilej că, în condiţiile confirmării ipotezei liniarităţii, pentru orice pereche de aporturi există un punct de echilibru în care ponderile numerice ale celor două influenţe sînt egale. Elementul care precumpăneşte înaintea punctului respectiv se vede, în schimb, surclasat după acest punct.

Acum, cel puţin în ceea ce priveşte productivitatea lexicală, ipoteza liniarităţii este confirmată. Putem, aşadar, să purcedem la o analiză, pe cît de concretă pe atît de exactă, a formelor de manifestare a „paradoxului lui A. de Cihac" în cîmpul vocabularului limbii române.

Afirmam mai sus că reversibilitatea raporturilor cantitative dintre două componente etimologice ale lexicului are caracter de lege. Sursa dominantă pînă la un anumit nivel de importanţă a cuvintelor devine



Scanned with OKEN Scanner

dominată de îndată ce se coboară sub acest nivel. În condiţiile unui vocabular numericeşte limitat, acest adevăr general trebuie însă înţeles numai ca o posibilitate, susceptibilă sau nu de actualizare. Teoretic, există într-adevăr un punct începînd cu care elementul maghiar precumpăneşte asupra celui latin moştenit, dar aceasta are loc doar într-un vocabular cuprinzînd cel puţin $10^{23}$ (altfel spus, o sută de mii de miliarde de miliarde!) de cuvinte. Practic, acceptînd limitele, şi aşa foarte largi, pînă la care se întinde tabelul 14.1., nici o influenţă, cu excepţia celei „franceze", nu ajunge să egaleze cantitativ fondul latin al limbii române.

În schimb, substratul este treptat depăşit numeric de toate celelalte aporturi. El cedează în faţa elementului vechi slav începînd cu $n = 15$, este ajuns din urmă şi întrecut de materialul lexical de origine onomatopeică la $n = 463$, de împrumuturile bulgăreşti la $n = 681$, de cele maghiare la $n = 826$, de cele „franceze" la $n = 1\,216$, de cele greceşti la $n = 1422$, de cele turceşti la $n = 1\,734$, de cele latine savante la $n = 2304$, de cele germane la $n = 2\,959$, de cele ucrainiene la $n = 4\,717$, de cele italieneşti la $n = 5\,027$, de cele ruseşti la $n = 6\,255$ şi de cele sîrbocroate la $n = 6\,875$. Rămîne totuşi remarcabil faptul că, de pildă, în ciuda dinamicii foarte active a aportului „francez", la nivelul celor mai productive 1 000 de unităţi, cuvintele comune cu albaneza devansează încă totalitatea neologismelor romanice („franceze" + italiene + latine savante).

În tot spaţiul explorat, elementul vechi slav este dominat de cel latin, însă domină, la rîndul său, toate celelalte aporturi, cu excepţia celui „francez", de care se lasă întrecut începînd cu $n = 3\,226$.

Acesta din urmă, deşi pornit în cursă în urma marii majorităţii a celorlalte influenţe, pe care nu le poate concura în profunzime, dar sfîrşeşte prin a le covîrşi numericeşte, depăşeşte succesiv: aportul sîrbocroat la $n = 488$, pe cel latin savant la $n = 535$, pe cel german la $n = 700$, pe cel turcesc la $n = 819$, pe cel grecesc la $n = 1\,121$, pe cel autohton la $n = 1\,216$, pe cel maghiar la $n = 1\,350$, pe cel bulgăresc la $n = 1\,448$, pe cel onomatopeic la $n = 1\,565$, pe cel vechi slav la $n = 3\,226$ şi, în fine, pe cel latin la $n = 5\,124$. O comportare, mutatis mutandis, asemănătoare o are elementul latin savant care, apărut abia la $n = 353$, egalează şi întrece pe rînd: aportul sîrbocroat la $n = 384$, pe cel german la $n = 1\,343$, pe cel autohton la $n = 2\,304$, pe cel maghiar la $n = 4\,793$, pe cel grecesc la $n = 5\,075$, pe cel bulgăresc la $n = 5\,910$, pe cel onomatopeic la $n = 6\,940$. Împrumuturile turceşti nu acced nici ele la straturi prea adînci ale vocabularului (primul dintre ele ocupă de-abia poziţia a 276-a în ierarhia productivităţii), dar densitatea lor sporită în zonele marginale ale limbii le permite să depăşească treptat: elementul german la



Scanned with OKEN Scanner

$n$ = 394, pe cel autohton la $n$ = 1 734, pe cel grecesc la $n$ = 2 463, pe cel maghiar la $n$ = 2 995, pe cel bulgăresc la $n$ = 3 661 şi pe cel onomatopeic la $n$ = 4 397. Un element activ la periferia vocabularului este cel italian care, debutînd modest (prima ocurenţă la $n$ = 613!), reuşeşte să treacă de cel rusesc la $n$ = 1 650, de cel sîrbocroat la $n$ = 3 280, de cel ucrainian la n = 5880, ba chiar şi de cel german la $n$ = 14 810 şi de cel maghiar la $n$ = 18 029! Alte inversări de poziţii mai au loc între aportul ucrainian şi cel sîrbocroat la $n$ = 1 041, între cel maghiar şi cel grecesc la $n$ = 4 232, între cel sîrbocroat şi cel rusesc la $n$ = 5 190 şi între cel bulgăresc şi cel grecesc la $n$ = 8 220.

Una dintre consecinţele acestei lungi serii de permutări şi remanieri este variabilitatea de la un „cerc" lexical la altul a ponderii relative a fiecărei componente etimologice. Întrebarea „cum se repartizează procentual pe origini cuvintele limbii române" la care s-au dat de-a lungul timpului răspunsuri diferite, adesea contradictorii, încetează astfel de a mai putea fi considerată drept o problemă bine pusă. Mai mult decît atît: apare posibilitatea manipulării procentelor în scopul supradimensionării rolului anumitor influenţe *fără a impieta totuşi asupra corectitudinii calculului!* Afirmaţia pare paradoxală, dar ea constituie un corolar firesc al observaţiilor de mai sus. Un simplu exemplu va lămuri lucrurile.

Să presupunem că ne interesează poziţia împrumuturilor vechi slave în ansamblul vocabularului limbii române. Cu datele din tabelul 14.1., e uşor de calculat procentul reprezentat de cuvintele de origine slavă veche din totalul $n$ de cuvinte înscris în prima coloană a tabelului. Iată cîteva dintre aceste valori:

*Tabelul 14.2*

| TOTAL CUVINTE | 10 | 50 | 100 | 200 | 400 | 600 | 800 | 1 000 |
|---|---|---|---|---|---|---|---|---|
| Cuvinte slave (%) | 10,0 | 12,0 | 13,0 | 14,50 | 16,25 | 17,17 | 17,38 | 17,50 |

| Total cuvinte | 1 100 | 1 200 | 1 400 | 1 600 | 1 800 | 2 000 | 5 000 | 10 000 |
|---|---|---|---|---|---|---|---|---|
| Cuvinte slave (%) | 17,55 | 17,42 | 17,29 | 17,00 | 16,72 | 16,45 | 12,30 | 8,72 |

Se observă că ponderea cuvintelor slave creşte monoton, atingând un maximum în jurul valorii $n$ = 1 100, după care începe să scadă (în principal ca urmare a înmulţirii vertiginoase a împrumuturilor romanice). Aşa se face că termenii slavi reprezintă 12% din cele mai productive 50 de cuvinte româneşti şi exact tot 12% (636) din totalul celor mai productive 5300 de cuvinte. Există însă o zonă mai restrînsă, situată între $n$ = 530 şi $n$ = 1 600, în cuprinsul căreia cota-parte reprezentată de elementul slav depăşeşte 17% din numărul total de cuvinte. Apexul acestei regiuni e situat în jurul valorii $n$ = 1 100. Aici



Scanned with OKEN Scanner

procentul unităţilor de origine slavă atinge cifra maximă: 17,55%. Aşadar, presupunând că ne-am propune să prezentăm componenta slavă a limbii române într-o lumină cît mai favorabilă, scopul nostru ar fi atins dacă ne-am axa analiza pe un vocabular reprezentativ de cca. 1 100 de cuvinte. Să notăm, în acest context, că variantele de fond principal lexical propuse de Al. Graur cuprind între 965 şi 1 419 cuvinte şi că cea recomandată drept „cea mai sigură şi demnă de încredere" conţine 1 181.

Într-un mod asemănător se stabileşte că un vocabular de cca. 2 000 de cuvinte este de natură să prezinte înrîurirea maghiară în postura ei cea mai avantajoasă, în timp ce unul de cca. 10 000 de cuvinte flatează influenţa turcească etc. Practic, cu excepţia fondului latin moştenit, a cărui pondere scade continuu odată cu extinderea domeniului lexical, şi cu cea a elementului „francez", supus unei mişcări exact contrare, pentru toate celelalte aporturi există cîte un punct privilegiat unde componenta respectivă apare potenţată.

Variaţiile discutate mai sus trebuie puse în legătură cu variabilitatea, de la o regiune a vocabularului la alta, a densităţii termenilor de o aceeaşi origine. Diferenţierea cu ajutorul caracterelor tipografice („de rînd" pentru cuvintele latine, cursive pentru celelalte) a unităţilor de obîrşie latină faţă de cele de alte origini, efecutată în tabelul 9.1., a pus în lumină modul cum, pe măsură ce ne îndreptăm către periferia vocabularului, cele dintîi cedează pasul în favoarea celor din urmă. Odată cu creşterea rangului, în şirul cuvintelor ordonate după un anumit criteriu de importanţă (care în cazul de faţă este productivitatea lexicală, dar ar putea fi prea bine şi un altul) cuvintele de diferite provenienţe se îndesesc sau se răresc potrivit unor legi de variaţie ce pot fi deduse lesne din ecuaţiile modelului $M_p$.

Ştim că legătura dintre numărul $n_i$ al cuvintelor de origine $i$ şi numărul $n_j$ al cuvintelor de origine $j$ este exprimată de o relaţie algebrică de forma:

$$ln\ n_i = a_{ij} ln\ n_j - b_{ij} \qquad (14.1)$$

care, delogaritmată, se mai poate scrie:

$$n_i = n_j^{a_{ij}} / e^{b_{ij}}$$

Densitatea relativă a cuvintelor de origine $i$ în raport cu cele de origine $j$ va fi dată de derivata:

$$\frac{dn_i}{dn_j} = \frac{a_{ij}}{e^{b_{ij}}} n_j^{a_{ij}-1} \qquad (14.2)$$

În şirul ordonat al cuvintelor, termenii aparţinînd celor două clase sînt egal de denşi în punctul $dn_i/dn_j = 1$, unde tangenta curbei face

cu axele de coordonate un unghi de 45°. Din forma logaritmată a expresiei (14.2) se deduce:

$$ln\left(\frac{dn_i}{dn_j}\right) = ln\, a_{ij} - b_{ij} + (a_{ij} - 1)\, ln\, n_j = 0$$

de unde:

$$ln\, n_j = \frac{b_{ij} - ln\, a_{ij}}{a_{ij} - 1}. \qquad (14.3)$$

formulă cu ajutorul căreia putem determina punctul de egală densitate a materialului lexical de origine *i* şi, respectiv, *j*. Să o utilizăm, de exemplu, pentru identificarea acelei zone a vocabularului ordonat după productivitate în care neologismele „franceze" ajung la fel de dense ca şi cuvintele de origine latină.

Forma particulară a expresiei (14.1) pentru relaţia „francez"-latin (v. paragraful 13) este:

$$ln\, n_{fr} = 4{,}748\ ln\, n_{lat} - 26{,}786$$

Avem deci $a_{ij} = 4{,}748$ şi $b_{ij} = 26{,}786$, valori care, introduse în formula (14.3), furnizează $ln\, n_{lat} = 6{,}731$, adică $n_{lat} = 823$, Punctul căutat se află aşadar după ocurenţa a 823 de cuvinte latine, ceea ce corespunde rangului $n = 2\,130$. Prin urmare, faptul deja cunoscut că, începînd cu $n = 5\,124$, elementul „francez" îl depăşeşte cantitativ pe cel latin se datoreşte densităţii superioare pe care cel dintîi a reuşit să o realizeze încă din punctul $n = 2\,130$. În general, fiecare punct de echilibru numeric dintre două componente etimologice diferite este precedat de o răsturnare a raportului dintre densităţi, fenomen care are loc într-o „secţiune" ce poate fi determinată cu ajutorul procedurii de mai sus. Cititorul este acum în măsură să calculeze poziţia acestor „secţiuni" pentru orice pereche de influenţe care îi suscită interesul, dar trebuie să aibă în vedere că forma concretă a relaţiei (14.1) nu i-a fost dată explicit pentru toate cuplurile de aporturi, motivul fiind economia de spaţiu, căci era vorba de vreo 200 de expresii diferite ce ar fi încărcat inutil lucrarea. Într-adevăr, parcurgerea etapei respective nici nu este strict necesară, dacă se ţine seama de legătura dintre parametrii $a_{ij}$, $b_{ij}$ şi constantele $m_i$, $m_j$, $k_i$, $k_j$ din ecuaţiile dreptelor caracteristice:

$$ln\, n_i = k_i - m_i d \quad \text{şi} \quad ln\, n_j = k_j - m_j d$$

particularizate pentru toate componentele etimologice principale ale vocabularului în capitolul 10. Această legătură este, potrivit ecuaţiei (13.4):

$$a_{ij} = \frac{m_i}{m_j}, \quad b_{ij} = \frac{m_i}{m_j} k_j - k_i. \qquad (14.4)$$



Scanned with OKEN Scanner

Introducînd (14.4) în (14.3), rezultă că secţiunea unde densităţile se egalizează este situată în punctul unde:

$$\ln n_i = \frac{m_i k_j - m_j (k_i - \ln m_j + \ln m_i)}{m_i - m_j}. \qquad (14.5)$$

Cu titlul de exemplu, să calculăm poziţia punctului de egală densitate a împrumuturilor bulgăreşti şi turceşti. Din relaţiile (10.6.2) şi (10.9) extragem: $m_{bg} = 0{,}384$, $m_{tc} = 0{,}690$, $k_{bg} = 5{,}597$, $k_{tc} = 6{,}232$. Cu aceste valori formula (14.5) devine:

$$\ln n_{bg} = \frac{0{,}384 \cdot 6{,}234 - 0{,}690(5{,}597 + 0{,}371 - 0{,}957)}{0{,}384 - 0{,}690} = 3{,}476$$

de unde $n_{bg} = 32$. Consultînd acum tabelul 14.1., vom conchide că în jurul rangului $n = 980$ materialul lexical de origine turcească atinge o densitate egală cu a celui provenit din limba bulgară.

## 15. Adecvarea modelului $M_p$ la realitatea vocabularului limbii române

Pentru a nu da loc la confuzii, să reamintim faptul că întreaga analiză întreprinsă mai sus a avut drept obiect nu realitatea concretă a vocabularului limbii române, ci un construct formal abstras din această realitate[1]: modelul $M_p$.

[1] De fapt, o atare subliniere apare ca necesară numai într-un domeniu încă insuficient matematizat, ea fiind, de regulă, trecută cu vederea în disciplinele care recurg în mod curent la modelarea matematică a fenomenelor. Un astronom poate anunţa, fără teama de a fi greşit înţeles, că a calculat traiectoria unei comete, cînd, în realitate, calculele sale se referă nu la corpul ceresc respectiv, ci la un punct material de masă egală cu a acestuia, obiectul studiului său constînd, în ultimă analiză, în modelul abstract construit cu ajutorul ecuaţiilor mecanicii generale, care simulează mişcarea cometei. O practică îndelungată a modelării matematice îl dispensează de a mai recurge la asemenea precizări „elementare", care însă pentru lingvist, slujitor al unei ştiinţe mult mai puţin evoluate sub acest aspect, sînt, cel puţin pentru moment, indispensabile.



Scanned with OKEN Scanner

Valoarea deducţiilor noastre pentru cunoaşterea lingvistică depinde, în mod hotărîtor, de fidelitatea reflectării în model a trăsăturilor pertinente ale obiectului modelat. De o conformitate totală nu poate fi, evident, vorba, cu atît mai mult cu cît, indiferent ce am crede despre caracterul sistematic al vocabularului, un ansamblu atît de complex nu se va putea supune niciodată unor reguli stricte, decît cu preţul lăsării pe dinafară a unui număr considerabil de excepţii.

Un exemplu care sare în ochi: potrivit modelului $M_p$, cel mai productiv cuvînt „comun cu albaneza" ar trebui să ocupe poziţia a cincea în ierarhia celor mai productive cuvinte româneşti. Tabelul 9.1 ni-l prezintă însă pe autohtonul *moş* drept cuvîntul care a generat cele mai multe derivate. În acest punct, modelul nu coincide aşadar cu realitatea modelată, dar neconcordanţa nu anulează valoarea modelului, deoarece majoritatea zdrobitoare a termenilor comuni cu albaneza se conformează într-adevăr aproape perfect ecuaţiilor stabilite, cuvîntul *moş* prezentîndu-se drept una dintre acele excepţii care nu fac decît să confirme regula[1].

Tabelul 15 ne oferă o imagine cuprinzătoare a raporturilor dintre modelul $M_p$ şi vocabularul limbii române. S-a arătat în capitolul 9 că ordonarea cuvintelor după productivitate determină în corpusul constituit de totalitatea termenilor nederivaţi cuprinşi în Dicţionarul Limbii Române Moderne un număr de 26 de clase distincte (v. tabelul 9.1). Potrivit acestei clasificări, 11 cuvinte au generat cel puţin 23 de derivate, 15 cuvinte — cel puţin 22 de derivate, 20 — cel puţin 21, 26 — cel puţin 20, 33 — cel puţin 19,.. şi 6 491 de cuvinte — cel puţin un derivat. Repartiţia pe origini a membrilor fiecărei clase a făcut obiectul analizei din capitolul 10. Tabelul 15 alătură această distribuţie celei pe care ne-o furnizează modelul $M_p$, permiţînd o comparare directă a rezultatelor calculului cu situaţia reală din D.L.R.M. şi evaluarea pe această bază a gradului de adecvare a modelului. Comentariile sînt, în general, de prisos, întrucît cifrele vorbesc de la sine. Mare parte dintre abaterile, de altfel puţin importante, pe care ni le relevă tabelul 15, au fost deja discutate cu prilejul analizei separate, pe componente etimologice, a lexicului românesc. Astfel, de exemplu, excepţional de productivul cuvînt *pic* este responsabil de

---

[1] Riguros vorbind, nici nu avem dreptul să pretindem unui element individual să se supună întocmai unei legi cu caracter statistic. Posibila sa abatere de la regula valabilă „în medie" este cum nu se poate mai naturală, ceea ce nu înseamnă încă că, în cazul depăşirii anumitor limite, ecartul nu poate îmbrăca o semnificaţie deosebită.



Scanned with OKEN Scanner

*Tabelul 15*

| TOTAL CUVINTE | | LAT. | cf. ALB. | SLAVE | INC. | ONOM. | BG. & SCR | BG. | MAGH. | NEC. | N.S.M | GR. | GERM | TC. | SCR. | LAT. SAV. | UCR. | „FR." | ROM. (LAT.) | RUS. | ITAL. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | calculat | 9 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 9 | 1 | — | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 15 | calculat | 13 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 13 | 1 | — | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 20 | calculat | 17 | 1 | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 17 | 2 | — | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 26 | calculat | 21 | 2 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 21 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 33 | calculat | 27 | 2 | 3 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 26 | 2 | 3 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 41 | calculat | 32 | 2 | 4 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 31 | 2 | 4 | 1 | 1 | — | 1 | — | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 50 | calculat | 38 | 3 | 6 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 37 | 2 | 6 | 1 | 1 | — | 1 | — | 2 | — | — | — | — | — | — | — | — | — | — | — |
| 62 | calculat | 47 | 3 | 7 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 45 | 2 | 9 | 1 | 1 | — | 1 | — | 3 | — | — | — | — | — | — | — | — | — | — | — |
| 82 | calculat | 60 | 4 | 11 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| | real | 57 | 3 | 12 | 1 | 1 | 2 | 1 | — | 5 | — | — | — | — | — | — | — | — | — | — | — |
| 100 | calculat | 72 | 5 | 13 | 3 | 2 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| | real | 70 | 4 | 13 | 2 | 1 | 3 | 2 | — | 5 | — | — | — | — | — | — | — | — | — | — | — |
| 126 | calculat | 88 | 6 | 17 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| | real | 87 | 5 | 17 | 2 | 1 | 3 | 3 | 1 | 5 | 2 | — | — | — | — | — | — | — | — | — | — |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 162 | calculat | 111 | 7 | 23 | 5 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| | real | 111 | 7 | 21 | 4 | 1 | 3 | 4 | 1 | 6 | 2 | 1 | — | 1 | — | — | — | — | — | — | — |
| 212 | calculat | 142 | 9 | 32 | 7 | 5 | 4 | 4 | 4 | 3 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| | real | 142 | 7 | 26 | 8 | 4 | 4 | 4 | 3 | 7 | 3 | 1 | 1 | 2 | — | — | — | — | — | — | — |
| 286 | calculat | 182 | 11 | 45 | 11 | 8 | 6 | 6 | 6 | 5 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — |
| | real | 187 | 8 | 37 | 11 | 8 | 5 | 6 | 3 | 8 | 3 | 3 | 1 | 4 | 1 | — | — | — | 1 | — | — |
| 375 | calculat | 231 | 13 | 61 | 15 | 12 | 8 | 9 | 8 | 9 | 2 | 3 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| | real | 231 | 13 | 56 | 13 | 12 | 6 | 8 | 6 | 11 | 3 | 5 | 1 | 7 | 1 | — | — | — | 2 | — | — |
| 479 | calculat | 290 | 16 | 84 | 22 | 17 | 12 | 13 | 12 | 14 | 3 | 6 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| | real | 286 | 15 | 77 | 23 | 16 | 11 | 11 | 14 | 18 | 4 | 6 | 1 | 7 | 1 | 2 | — | 1 | 3 | — | 1 |
| 659 | calculat | 362 | 20 | 114 | 30 | 24 | 16 | 19 | 17 | 22 | 5 | 9 | 4 | 4 | 2 | 3 | 2 | 3 | 1 | 1 | 1 |
| | real | 374 | 19 | 102 | 28 | 19 | 16 | 16 | 17 | 27 | 4 | 11 | 1 | 7 | 3 | 5 | 1 | 5 | 3 | — | 1 |
| 929 | calculat | 471 | 25 | 162 | 45 | 36 | 24 | 30 | 27 | 38 | 7 | 16 | 7 | 10 | 4 | 6 | 4 | 11 | 2 | 2 | 2 |
| | real | 505 | 32 | 149 | 38 | 36 | 22 | 31 | 20 | 35 | 6 | 14 | 5 | 9 | 6 | 6 | 1 | 9 | 4 | — | 1 |
| 1250 | calculat | 583 | 30 | 217 | 62 | 51 | 32 | 43 | 38 | 60 | 10 | 26 | 12 | 19 | 6 | 12 | 7 | 32 | 3 | 3 | 4 |
| | real | 621 | 32 | 214 | 58 | 51 | 32 | 46 | 31 | 50 | 7 | 20 | 12 | 16 | 7 | 10 | 7 | 24 | 7 | 2 | 3 |
| 1728 | calculat | 725 | 35 | 291 | 85 | 72 | 44 | 61 | 54 | 93 | 14 | 42 | 21 | 35 | 11 | 24 | 12 | 88 | 7 | 7 | 7 |
| | real | 769 | 35 | 305 | 79 | 74 | 40 | 66 | 49 | 80 | 11 | 40 | 19 | 33 | 10 | 18 | 14 | 63 | 9 | 7 | 7 |
| 2418 | calculat | 884 | 42 | 382 | 115 | 98 | 58 | 85 | 75 | 141 | 19 | 65 | 35 | 64 | 19 | 45 | 22 | 227 | 13 | 14 | 15 |
| | real | 924 | 39 | 412 | 109 | 104 | 53 | 91 | 79 | 108 | 12 | 66 | 33 | 65 | 20 | 37 | 20 | 203 | 13 | 15 | 15 |
| 3775 | calculat | 1110 | 51 | 520 | 162 | 140 | 80 | 125 | 110 | 225 | 27 | 106 | 62 | 127 | 33 | 93 | 42 | 670 | 26 | 30 | 36 |
| | real | 1105 | 51 | 535 | 159 | 140 | 80 | 112 | 112 | 165 | 19 | 106 | 53 | 124 | 33 | 86 | 37 | 763 | 25 | 30 | 40 |
| 6491 | calculat | 1394 | 62 | 708 | 229 | 201 | 111 | 183 | 160 | 361 | 39 | 175 | 111 | 252 | 60 | 193 | 80 | 1974 | 52 | 64 | 82 |
| | real | 1328 | 62 | 676 | 251 | 191 | 102 | 164 | 184 | 310 | 31 | 185 | 132 | 260 | 60 | 184 | 69 | 2101 | 54 | 67 | 80 |



Scanned with OKEN Scanner

mica inadvertenţă a priorităţii componentei onomatopeice faţă de elementul vechi slav (dintre cele mai productive 11 cuvinte, 9 sînt într-adevăr latine şi unul comun cu albaneza, adică exact cum rezultă din model, dar al 11-lea nu e slav, ci onomatopeic; abaterea se menţine şi pentru primele 15 cuvinte, între care, de această dată, 13 sînt latine). O altă neconcordanţă, probabil pasageră, dintre realitate şi model priveşte superioritatea în primele clase lexicale a numărului efectiv de cuvinte cu etimologie necunoscută în raport cu numărul scontat pornind de la ecuaţiile modelului. Dintre cele mai prolifice 82 de cuvinte, 5 sînt de provenienţă necunoscută, deşi potrivit calculelor nu ar fi trebuit să existe decît un singur cuvînt de acest fel. Originea autohtonă (foarte probabilă pentru *pînză* şi *băiat*) sau slavă (posibilă pentru *vorbă* ori *pungă*) a celorlalte 4 ar explica pe deplin decalajul observat.

În general, numărul real de cuvinte de diferite origini înregistrează oscilaţii, de amplitudine nu prea mare, în jurul valorii scontate. După ce au fost mai puţine decît prevedea modelul în clasele cu $n = 11$, 15, 20, 26, cuvintele slave se situează la nivelul calculat pentru $n = 33$, 41, 50, cresc uşor peste acesta pentru $n = 62$, 82, revin la situaţia de echilibru pentru $n = 100$, 126, scad sub valoarea aşteptată în zona $n = 162$, 212, 286, 375, 497, 659, 929, 1250, spre a o depăşi din nou pe sectorul $n = 1\,728$, 2 418, 3 775 şi a coborî în final sub ea pentru $n = 6\,491$. Asemenea fluctuaţii constituie o regulă, explicabilă prin procedura folosită la trasarea dreptelor caracteristice, metoda celor mai mici pătrate realizînd o mediere între punctele „experimentale" a căror dispunere nu este perfect coliniară, ceea ce face ca ele să se situeze alternativ deasupra şi dedesubtul dreptei în cauză.

Per total, examinarea tabelului 15 este liniştitoare. Modelul se dovedeşte apt să reproducă, cu o precizie absolut remarcabilă, structura pe origini şi niveluri de productivitate a vocabularului limbii române. Constatarea acestei adecvări validează implicit atît rezultatele şi concluziile obţinute aici din analiza modelului $M_p$, cît şi, anticipat, pe cele ce vor mai putea fi evidenţiate prin cercetări viitoare.

## 16. Criteriul polisemiei. Distribuţia pe origini. Ecuaţiile modelului $M_S$

Dintre parametrii revelatori pentru importanţa cuvintelor, ne-am ocupat pînă în prezent numai de productivitatea lexicală. Poziţia unui cuvînt în sistemul vocabularului depinde însă nu mai puţin şi de numărul de înţelesuri distincte pe care acesta le posedă. Cunoaşterea



Scanned with OKEN Scanner

structurii lexicului românesc ar fi în consecinţă lacunară dacă modelului productivităţii $M_p$ nu i-am alătura un model $M_s$ al polisemiei lexicale.

Metodologia elaborării acestui nou model este, în linii generale, aceeaşi cu cea aplicată anterior: se procedează mai întîi la ordonarea corpusului lexical (identic cu cel din cazul precedent) după numărul de sensuri şi provenienţă şi se încearcă să se identifice expresia matematică a legităţii ce guvernează clasificarea respectivă.

După numărul de înţelesuri consemnate în dicţionar, cuvintele din DLRM se grupează în 31 de clase. Enumerăm în tabelul 16.1 termenii ce aparţin primelor 26 dintre ele, reliefînd prin folosirea unor caractere tipografice distincte supremaţia fondului latin în zona rangurilor superioare şi creşterea progresivă a numărului elementelor non-latine (notate cursiv) pe măsură scăderii gradului de polisemie a cuvintelor.

*Tabelul 16.1*

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| 1 | 1 | de (prep.) | 47 |
| 2 | 2 | ţine | 46 |
| 3 | 3 | face | 45 |
| 4 | 4 | da (vb.) | 36 |
| 5 | 5 | pe | 31 |
| 6 | 6—7 | trage, trece | 30 |
| 7 | 8 | bun | 28 |
| 8 | 9 | la | 24 |
| 9 | 10 | lua | 23 |
| 10 | 11—12 | prinde, rău | 22 |
| 11 | 13—15 | drept, ridica, scoate | 21 |
| 12 | 16 | bate | 20 |
| 13 | 17—18 | fi, tot | 19 |
| 14 | 19—20 | după, tăia | 18 |
| 15 | 21 | ochi | 17 |
| 16 | 22—28 | cît, greu, în, lega, legătură, sta, şi | 16 |
| 17 | 29—30 | pune, tare | 15 |
| 18 | 31—36 | bine, cum, *de* (conj.), limbă, strica, umbla | 14 |
| 19 | 37—42 | avea, din, duce, faţă, intra, *linie* | 13 |
| 20 | 43—51 | cap, cădea, cerc, cu, întinde, pentru, strîns, vedea, veni | 12 |



Scanned with OKEN Scanner

*Tabelul 16.1. (continuare)*

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| 21 | 52—66 | căuta, ce, coardă, *colţ,* cruce, curge, cuvînt, pierde, pui, *punct,* roşu, *spirit,* striga, şti, *vorbă* . . . . . | 11 |
| 22 | 67—82 | apuca, aşeza, bătaie, corn, curat, deschide, floare, inimă, întoarce, întru, mai, merge, piatră, scaun, sub, strînge. . . | 10 |
| 23 | 83—110 | ajunge, arde, *băga, chip, corp,* cuprinde, dacă, dintru, juca, lăsa, mic, *mos,* nebun, nu, numai, parte, peste, petrece, picior, purta, răspunde, să, sări, *sfînt, străin,* tărie, ureche, *val* | 9 |
| 24 | 111—159 | ascuţit, aşa, capră, carte, că, cădere, călca, *centru,* cerca, *clasă,* coadă, *coroană, cos,* cunoaşte, *curent,* deşert, dulce, foc, *forţă,* fus, guşă, închis, întins, legat, *lovi,* lume, lung, mare, măsură, mijloc, mişcare, mult, *nervură,* nod, număr, pană, pătul, *pica, potrivi, prost,* puţin, scris, semn, sparge, strat, stricat, subţire, *ton,* vărsa. . . . . | 8 |
| 25 | 160—235 | aduna, alege, aripă, arunca, aspru, bunătate, *butuc,* ca, căutare, *cinste, cîrlig,* coace, *coborî, cumpăni, direcţie,* domn, drege, *element, funcţie, galerie, gînd,* greutate, gură, *hop,* împlini, închide, încurca, întări, *liber,* loc, lumina, lumină, *moţ,* negru, nostru, *notă,* om, *pat,* păpuşă, pişca, pînă, *platformă, pod, pomeni, prag,* prin, *raport, rază, răsuci,* rupe, scară, *seamă,* sec, *secret,* sîn, *slobod, sondă,* steluţă, stinge, stins, *stradă,* suflet, suna, şedea, şterge, *tablou, talpă,* tată, *termen, treabă,* turna, umbră, unde, urca, urma, *valoare* . . | 7 |
| 26 | 236—369 | ac, adînc, aduce, aprinde, aprins, asupra, atinge, *atîrna, bară, bază, bloc, bombă,* bucată, *burduf, calotă,* casă, cărbune, cel cere, ceva, cheie, cînd, *coajă, coloană,* crăpa, creşte, cuib, curăţa, *cursă, dar, deosebit, derivat,* desface, *domol,* dor, facere, fierbe, *fila, focar, fond, formă, gîndi, gol,* gros, gust, ieşi, *împleti,* înălţa, încă, încărcat, închina, însemna, înşira, *jgheab,* jos, joc, *jurnal,* luminat, meu, mişca, mînca, mort, muiat, *neam,* nou, *ocol, ordin,* pace, pămînt, pierdut, *piesă, pînză, plan,* plin, *pop, poştă,* precum, *prezenta, profil,* purice, *pustiu,* putere, *radical,* răsări, *reprezenta, reproduce,* reveni, *revoluţie, rînd,* sarcină, *sălas,* sămînţă, scăpa, scrie,scula, *sigur, simplu,* sînge, *socoti, sonor,* spune, *stil, stîrni,* strecura, strigare, strigăt, strîmb, strînsoare, sufla, sus, *tablă, tacîm,* tăietură, timp, *titlu, topi, trata, trăi, trezi,* trimite, *trup,* ţară, *ţintă,* umple, urechelniţă, urmări, usca, vechi, verde, vînă, vostru, vrea, vulpe. . . . . | 6 |



Scanned with OKEN Scanner

Compararea tabelelor 9.1 şi 15.1 evidenţiază faptul că, sub raport semantic, elementul latin ocupă o poziţie chiar mai puternică decît pe planul productivităţii lexicale. Printre cele mai productive 82 de cuvinte româneşti, 57 sînt de origine latină, adică cca. 69,5%, în timp ce, din cele mai bogate în sensuri 82 de cuvinte, materialul moştenit reprezintă nu mai puţin de 76 de unităţi (92,7%). În plus, merită semnalat faptul că jumătate dintre cei 6 termeni rămaşi sînt de provenienţă latină savantă (e vorba de cuvintele *linie, punct,* şi *spirit*).

În legătură cu această din urmă constatare, este locul să observăm că, dacă numărul de derivate constituie un indiciu sigur al vechimii cuvintelor, nu acelaşi lucru se poate afirma şi despre numărul de înţelesuri. Împrumuturi recente pot prezenta un nivel ridicat de polisemie, avînd în vedere faptul că la pătrunderea lor în româneşte pot aduce cu ele o bună parte din zestrea semantică agonisită în limba din care provin. Nici unul dintre sensurile „titlu universitar", „autorizaţie dată unui inventator pentru exploatarea unui brevet de invenţie", „încălcare a limitelor bunei-cuviinţe" sau „abatere de la normele gramaticale sau prozodice" pe care le cunoaşte cuvîntul *licenţă* nu a luat naştere în româneşte, absolut toate sînt împrumutate. Acelaşi lucru se poate afirma şi despre majoritatea înţelesurilor celor 3 cuvinte latine savante pomenite mai sus. Concluzia ce se desprinde de aici este că nu se pune problema să legăm cronologia diverselor aporturi lexicale de statistica sensurilor, aşa cum am corelat-o anterior cu statistica derivatelor. O atare tentativă ar avea poate unele şanse de succes în cazul că s-ar referi numai la sensurile noi, apărute pe teren românesc[1], dar pentru precizarea acestora este necesară o analiză semantică extrem de laborioasă a întregului vocabular, operaţiune care, pe lîngă marele volum de muncă solicitat, mai necesită şi rezolvarea în prealabil a unor probleme delicate, cum ar fi precizarea regimului calculului semantic, fenomen cu statut ambiguu, întrucît vizează sensuri care, deşi au apărut în româneşte, se conformează totuşi unui model străin. Nu ne lăsăm ademeniţi în acest hăţiş, cu atît mai mult cu cît sîntem departe de a ne fi propus să rezolvăm toate problemele lexicului. Obiectivele noastre sînt cu mult mai modeste şi deliberat limitate la cîteva idei de bază, dintre care în paragraful de faţă o urmărim pe cea a exprimării în termeni matematici a raporturilor

[1] La fel am procedat, de altminteri, şi în cazul derivatelor, neluîndu-le în considerare pe cele ce nu s-au format în româneşte, ci au fost împrumutate ca atare din limba care ne-a furnizat şi termenul-matcă.



Scanned with OKEN Scanner

dintre originea cuvintelor şi partiţia mulţimii lor în clase de echivalenţă pe criteriul numărului de înţelesuri distincte. Clasificarea pe care o întreprindem are în vedere exclusiv aceşti doi parametri (provenienţa şi gradul de polisemie) şi, drept urmare, dihotomia anterior practicată între cuvintele-matcă şi derivatele lor încetează de a mai fi pertinentă. Aşa se explică prezenţa în tabelul 16.1 a unor cuvinte ca *strîns, nebun, tărie, ascuţit, cădere, închis, legat, mişcare, potrivi, scris, stricat, căutare, coborî, întări, răsuci, steluţă, stins, urma* etc. formaţi în româneşte de la termeni primari prezenţi *(bun, căuta, lega, tare)* sau nu *(ascuţi, potrivă, suci)* printre elementele tabelului.

Începem, cum e şi firesc, cu examinarea moştenirii latine. Dat fiind că pînă la nivelul de polisemie 6 inclusiv, cuvintele de origine latină au fost deja clasificate (v. termenii tipăriţi cu caractere „de rînd" în tabelul 16.1), enumerăm mai jos numai unităţile din clasele 27, 28 şi 29:

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| 27 | 255—401 | abate, acoperi, adunare, alerga, an, apă, arătа, arsură, aşterne, baie, bătut, braţ, broască, cale, care, căluş, cerere, chema, chemare, chiar, cîmp, copt, cot, crede, credinţă, cuget, cupă, cuprins, curte, despărţit, desprinde, destul, deşira, dulceaţă, eu, făcut, ferecat, foaie, frînt, frumos, fugi, fum, găinuşă, gras, intrare, înainta, înălţime, încărca, încet, încheiat, încrucişa, încrucişat, încurcat, îndesat, îndoi, îndoit, îndrepta, înghiţi, îngâna, înţelege, învăţa, învăţat, larg, lămuri, lege, limpede, limpezi, lingură, lucra, lucru, lună, masă, măsura, mărunt, minte, mînă, moale, moarte, muia, muscă, nume, orb, palmă, pas, păr, pătrunzător, pereche, piept, pieptar, pîntece, plînge, poartă, pradă, pumn, punte, putea, puternic, râde, rar, ras, rădăcină, rămăşiţă, rămîne, răsărit, răscruce, roade, roată, rost, rosti, rugină, rupt, sălbatic, scădea, schimba, scurge, scutura, seca, singur, sîngerat, soră, spic, spin, spre, spurcat, stare, stea, sui, tău, toca, tremurat, trunchi, tu, tuna, ţese, uita, umăr, unu, urzi, uscat, uşure, vedere, viu, vîntura, zăcea, zbura, zburător, zice . . . . . | 5 |
| 28 | 402—657 | abia, adînci, afla, afumat, alb, albitură, ales, amiază, anume, apleca, apoi, apropia, apus, arc, ars, aţă, auzi, ăsta, barbă, băşică, bătător, bătrîn, bea, berbec, boteza, brumat, brumă, bute, ca, căluşel, căpătîi, căpăţînă, căruţ, căscat, căşuna, căţel, căzătură, cep, cerceta, ceresc, cerne, ciur, cînta, cîştiga, coastă, contrazice, corb, cotitură, creştere, crud, crunt, | |



Scanned with OKEN Scanner

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| | | cufunda, culcuş, culege, cunoştinţă, cunună, curs, cusătură, dat, dres, descărca, desfunda, despica, desprindere, despuia, destinde, detunat, dezarma, dezlega, doar, dormi, des, drăcie, dreptate, fost, făţui, fereca, fir, frate, frecătură, frînge, fund, fuscel, ghiaţă, gusta, iarbă, iernatic, inel, împărat, împărtăşi, împinge, împlinit, înalt, înăspri, începea, încercare, încheia, îndărăt, îndulci, îneca, înecat, înfăşura, înfierat, înflori, înfundat, întărit, înteţi, întinsoare, întors, între, întreg, întunecat, învăţătură, învoi, jude, judeca, jug, lapte, latură, lemn, lepădа, lin, lîngă, lucrător, luminos, lumînare, lungi, lup, lupta, luptă, mamă, martor, mărime, măsurat, mătreaţă, merişor, mers, moliciune, morişcă, muchie, mura, mut, noi, numi, pată, păi, pătrat, pătrunde, pieliţă, pieptar, pieptiş, pietrărie, pişcătură, piuă, pîine, plumb, poamă, prădа, prescrie, preţ, prevedea, prinsoare, printre, răscumpăra, răsfrînge, răsturna, răsufla, răutate, răzbate, reînnoi, repede, repezi, retrage, reţea, revărsa, ridicat, român, rostui, roti, rotund, rug, rugă, ruginit, ruşine, sac, sau, săgeată, săgeta, sălta, săpa, săpat, sărat, săricică, său, schimb, schimbător, scoarţă, scriitor, scriptură, scrisoare, scurt, scut, simţi, simţire, sîngera, sîngeros, smulge, spart, spălător, spărgător, spînzura, stat, stoarce, strîmt, strîmtoare, substrat, subţia, suferi, suge, sul, sunător, sunet, supune, surpătură, şa, şuier, şuiera, tăcea, tăios, tîrziu, trăgător, trecut, tremura, tremurici, tulburare, turbat, ţîţă, umfla, umflat, umplutură, undă, unghie, urît, urmă, urs, uşor, varză, vecinătate, viaţă, viitor, viţel, vînăt, vînt, voi (pron.), voinţă, zăcătoare . . . . . | 4 |
| 29 | 658—1088 | acoperire, acrit, adevăr, adevărat, adia, adîncime, adormi, aer, afunda, aici, alinta, altminteri, alunar, aluneca, alunel, alunele, amar, amesteca, apăra, apărător, arămărie, arbore, ardere, arici, arma, asculta, asmuţi, astupa, aşezămînt, aştepta, atît, atrage, atunci, bărbie, bătătură, beţie, binişor, botez, brăţară, brîncă, bucătărie, bulbuc, bunic, buric, cal, car, carne, călător, călcare, căldare, cămaşă, căprior, căsca, căţea, cenuşar, cer, cercare, cerceluş, cercui, certa, cheotoare, chingă, ciumă, cîntător, coamă, coase, colea, consemna, coptură, cornut, coti, creastă, creier, crescut, cuceri, cufunda, cuibări, culca, culme, cumpăt, cuptor, curea, curelărie, curgător, cuvînta, datora, degera, deget, degeţel, descurca, desfăcut, desfăşura, desfereca, desferecat, desfundat, despărţi, | |



Scanned with OKEN Scanner

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| | | despărţire, despărţitură, despuiat, destupa, deştepta, detuna, dezlegat, dezvolta, dezvoltare, dinte, doagă, doamnă, dogărie, domnie, dori, el, făptură, făţuială, fecioară, fecior, fereastră, feri, fier, fiere, fire, foi, forfeca, fragă, frămînta, freca, frige, frîntură, frunte, frunzar, fugar, fulgera, funie, furcă, gălbinare, găleată, geană, grătar, grăunte, grosime, gustare, ierta, ieşire, îmbrăca, îmbrăcăminte, îmbrăţişa, îmbuca, împăna, împărăţie, împărţi, împiedica, împresura, împunge, înaintat, înainte, înapoia, încăleca, încălţa, începe, încerca, închega, încinge (lat. incendere), încinge (lat. incingere), înconjura, încorda, încordat, încredinţa, încreţi, îndatora, îndesa, îndura, înfăţişa, înfierbînta, înfiora, înflori, înfoia, înfoiat, înfunda, îngheţa, îngreuia,îngreuna, îngroşa, îngust, înmuia, înmuiat, înnegri, înnoi, înroşi, înscrie, însemnat, însoţi, însuşi, înşela, înşirui, întăritură, întărîta, întîmpla, întîrzia, întorsătură, întreba, întrece, întreţine, întuneca, întuneric, înţelege, învăţămînt, învăţător, înverzi, învoalbe, învoială, învoire, înzestra, judecată, lăcrămos, lămurit, leagăn, legăna, lemnar, lemnărie, limpezime, limpeziş, linge, linguriţă, liniştit, linte, lînă, luci, luciu, lucrare, luminare, luminiş, luminiţă, lunatic, luneca, margine, măduvă, măiestrie, mănuşă, mărgea, mărgini, mărire, mărunţiş, miez, mînă, minunat, minune, mistreţ, mînă, mînecă, moară, molîu, munte, murat, muri, muscar, mustaţă, muşca, muşuroi, muta, naşte, nebunie, neted, netezi, nor, noutate, nucă, număra, oală, ochi, olărie, ori, os, par, paşnic, păcat, păducel, păduche, pădurărie, păpuşoi, părere, păretar, părinte, păşi, pătură, peniţă, perete, perişor, peşte, piedică, pieptăna, pieptănuş, pieptene, pieri, pierit, pierzanie, pietriş, pietros, pietroşel, pieziş, pisa, pişcător, plăpînd, plumbui, plumbuit, port, portar, porumb, preda, prefaçe, prefira, prelucra, preţui, prevedere, prim, prins, priveghea, puică, pulpă, purtare, purtător, răceală, răci, răpi, rări, răsfăţ, răsfăţa, răsfira, răsfirat, răsuna, răşchira, răşinos, rătăci, răzbuna, recunoaşte, recunoscut, repeziş, rîde, rîu, rîura, rotunji, rotunjit, ruga, rugini, rumega, ruptură, sapă, sănătos, săpătură, sărit, scălda, scăldătoare, scărmăna, scriitură, scufunda, scurgere, scursoare, scursură, scurta, scuti, scuturătoare, scuturătură, se, searbăd, semănătură, seminţie, singuratic, singurătate, sînziană, soare, soartă, soaţă, sorbi, | |

| CLASA | NR. DE ORDINE | CUVINTE | NR. DE SENSURI |
|---|---|---|---|
| | | spate, spălat, spălătură, spărtură, speria, spinare, spinteca, sprînceană, spumă, spurca, staul, stătător, sterp, stîng, stîrpi, străbate, străfiga, străfulgera, strălucitor, strămuta, străveziu, stricăciune, strîmba, strînsură, subţiat, sumete, supăra, supărare, supus, surd, surpa, suspin, şedere, şir, ştergător, tămîia, tămîioară, tărăgăna, toacă, toartă, tragere, trăgătoare, treaptă, trei, tremur, trimitere, tufan, tufă, tulbura, tulbure, turmă, turtit, tuşina, ţesătură, ţesut, ţinut, ţinută, uda, umbri, udătură, umezeală, unge, unghi, unime, unit, unul, urgie, urmaş, ursoaică, urzeală, urzică, uşura, uşurinţă, vad, vas, vărsare, vărsat, văz, văzut, veşted, veşteji, vier, viespar, vin, vineţea, vineţeală, vinos, vintre, vioară, vioi, vis, visa, vită, viţelar, vîlţoare, vînzare, vîrtute, volbură, vultur, zău, zbate, zburătăci, zeamă, zgaibă, zi . . . . . | 3 |

Ne oprim aici cu menţionarea nominală a termenilor de origine latină, avînd în vedere că ultimele două clase, a 30-a şi a 31-a, cuprind împreună 7 744 de elemente (cuvinte-bază şi derivate), adică de peste 7 ori mai multe decît toate celelalte la un loc. Notînd cu $s$ numărul de sensuri şi cumulînd, după cum s-a procedat şi în cazul productivităţii lexicale, clasele de cuvinte ce posedă cel puţin $s$ sensuri, se obţine, pentru gama de valori $s = 1 \div 23$, situaţia din tabelul 16.2.

*Tabelul 16.2.*

| $s \geq$ | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv. | 8832 | 1088 | 657 | 401 | 254 | 177 | 132 | 95 | 76 | 60 | 49 |
| ln (Nr. cuv.) | 9,09 | 6,99 | 6,49 | 5,99 | 5,54 | 5,18 | 4,88 | 4,55 | 4,33 | 4,09 | 3,89 |

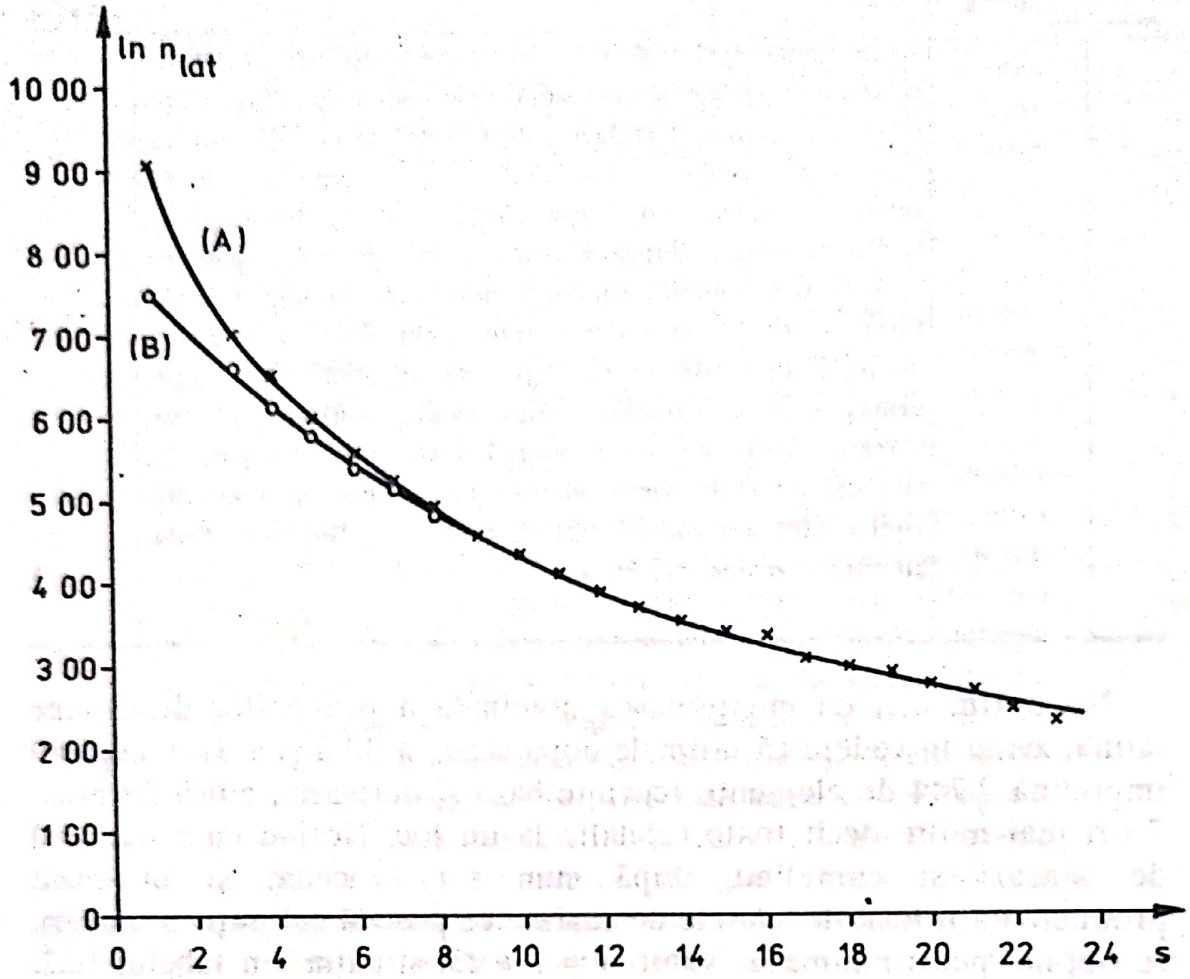
| $s \geq$ | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv. | 40 | 35 | 30 | 28 | 21 | 20 | 18 | 16 | 15 | 12 | 10 |
| ln (Nr. cuv.) | 4,69 | 3,56 | 3,40 | 3,33 | 3,04 | 3,00 | 2,89 | 2,77 | 2,71 | 2,48 | 2,30 |

Reprezentate într-un sistem rectangular de coordonate avînd în abscisă numărul de sensuri şi în ordonată logaritmul numărului de cuvinte (fig. 16.1), punctele corespunzătoare valorilor din tabel se



Scanned with OKEN Scanner

FIG. 16.1.




dispun de-a lungul unei curbe (A) cu concavitatea îndreptată în sus. Spre deosebire de cazul, anterior studiat, al producţiei de derivate, dependenţa în raport cu numărul de sensuri nu mai prezintă deci un caracter liniar. Curba respectivă debutează printr-o pantă coborîtoare foarte pronunţată, care scade rapid, tinzînd către o valoare relativ constantă. Graficul pare să posede aşadar o asimptotă înclinată. Pentru a doua jumătate a curbei, aproximarea cu o dreaptă ar mai fi deci posibilă, lucru de neconceput, în schimb, în ceea ce priveşte prima sa porţiune.

Prin luarea în considerare numai a cuvintelor-bază, primele 12 clase din tabelul 16.2., singurele care cuprind şi derivate, suferă modificări cu atît mai importante cu cît se referă la valori mai mici ale parametrului *s*. Noua curbă (B), trasată pe aceeaşi diagramă cu precedenta, se identifică cu aceasta începînd cu punctul de abscisă $s = 13$, ceea ce înseamnă că toate cuvintele moştenite din latineşte care au cel puţin 13 sensuri sînt termeni primari, nederivaţi. În schimb



Scanned with OKEN Scanner

ramura sa stîngă se îndepărtează sensibil (foarte mult, de fapt, dacă se ţine seama că reprezentarea din fig. 16.1. este logaritmică: în valori absolute, ordonata punctului de abscisă s = 1 este de aproape 5 ori mai mică pentru curba (B) decît pentru (A)!) de ramura corespunzătoare a lui (A), neliniaritatea relaţiei fiind aici mai puţin pregnantă decît în cazul dintîi. Coordonatele punctelor în care (B) diferă de (A) (menţionăm şi de această dată că graficele noastre au sens numai în punctele de abscisă întreagă, căci, la fel cum nu există cuvinte care să fi generat un număr fracţionar de derivate, nu există nici cuvinte care să posede un număr neîntreg de înţelesuri) sînt prezentate în tabelul 16.3.

*Tabelul 16.3.*

| $s \geq$ | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 1802 | 753 | 465 | 323 | 221 | 161 | 121 | 92 | 74 | 59 | 48 |
| ln (Nr. cuv.-bază) | 7,50 | 6,62 | 6,14 | 5,78 | 5,40 | 5,08 | 4,80 | 4,52 | 4,30 | 4,08 | 3,87 |

Căutarea unor expresii analitice pentru curbele (A) şi (B) este destul de dificilă şi, în plus, s-ar putea să nici nu fie strict necesară pentru elaborarea modelului $M_s$. Într-adevăr, din experienţa construirii omologului său *Mp*, ştim că în forma finală a ecuaţiilor singurii parametri independenţi care mai apar sînt numerele de cuvinte de diferite origini. Dependenţa în raport cu numărul de derivate nu mai este explicită, situaţie care ar putea fi valabilă şi în ceea ce priveşte numărul de sensuri ale cuvintelor. Să mai amînăm deci momentul exprimării analitice a raporturilor numerice, examinînd deocamdată cum se prezintă lucrurile pentru o altă sursă lexicală bogat reprezentată în mulţimea cuvintelor cu grad de polisemie ridicat: neologismele „franceze". Iată lista termenilor de origine „franceză" care figurează în DLRM cu cel puţin 3 sensuri distincte:

1. corp . . . . . 9 sensuri
2—7. centru, clasă, curent, forţă, nervură, ton 8 sensuri
8—19. direcţie, element, funcţie, galerie, liber, notă, platformă, raport, secret, sondă tablou, valoare . . . . . 7 sensuri
20—42. bară, bază, bloc, bombă, calotă, coloană, cursă, derivat, fila, fond, formă, jurnal, piesă, plan, prezenta, profil, radical,



Scanned with OKEN Scanner

reprezenta, revoluţie, simplu, sonor, stil, trata . . . . . 6 sensuri

43—93. acord, aparat, asimila, atac, ataca, bandă, cadru, capsulă, celulă, colonie, compoziţie, concentra, condiţie, cordon, damă, declara, familie, figură, fixa, formaţie, formulă, grad, graţie, intim, istoric, înclina, lojă, major, marcă, mască, ocupa, opera, operaţie, opoziţie, pozitiv, poziţie, priză, proiecta, real, registru, rest, rezistenţă, rula, sens, serios, sifon, solid, suspensie, susţine, teatru, volum . . . . . 5 sensuri

94—221. abataj, absolut, acţiune, altera, *alterat,* antrena, aplica, articol, autoritate, banc, baterie, bazin, bibliografie, birou, buclă, buletin, cantitate, capacitate, carton, ceda, circula, clasic, comandă, comisar, compliment, concurent, condamna, congregaţie, conştiinţă, conţinut, convertizor, corespondent, cotă, declina, depresiune, derivaţie, *desrămat,* diapazon, disc, dispune, dragon, echilibru, economic, efemeridă, emisiune, etaj, evacua, evoluţie, executa, explozie, extrem, fantastic, fantezie, filozofie, filtru, forma, fracţiune, front, garnitură, gaz, general, generator, grav, idee, imagine, inferior, interior, lamă, libertate, licenţă, liră, mandat, manevră, maşină, material, membru, minister, motiv, natural, natură, observa, observaţie, onoare, organ, parc, pasiv, perioadă, personalitate, pilă, piramidă, pistă, pleiadă, popular, post, practică, producţie, proprietate, rampă, raporta, recepţie, redresor, reforma, reformă, regim, *reprezentare,* reversibil, rezervă, rozetă, segment, serie, servi, serviciu, sferă, societate, stabilizator, staţiune, stigmat, talon, tensiune, transmite, tren, tub, unitate, universal, versiune, vibrator, voluntar, vulgar . . . . . 4 sensuri



Scanned with OKEN Scanner

222—697. aberaţie, accent, accident, achita, acorda, activa, aferent, afinitate, agent, agricol, alcool, *aliat,* ambala, amfiteatru, angaja, antenă, antic, apel, apendice, aranja, arenă, armură, articula, *articulat,* artificial, asociaţie, atenţie, atracţie, avansa, avantaj, bac, balansa, balansier, balon, bandaj, bibliotecă, bilateral, bituma, bonetă, bridă, brigadă, brigadier, broşă, buchet, bufet, bulă, bulb, cabinet, cală, calibra, calibru, califica, calitate, canal, canevas, canton, capital (subst.), capital (adj.), caracter, caracteristică, careu, centra, centrală, centură, cenzor, cerebral, cheson, ciclu, cilindru, circ, circuit, circulaţie, cita, client, clişeu, colaborator, colecta, colectiv, comandament, combate, comisariat, companie, compresor, comun, comună, concluzie, concura, concurs, conductor, conferinţă, confirma, congres, *conjugare,* conjunctiv, consideraţie, consilier, consolă, construcţie, construi, consultaţie, consuma, conta, contingent, control, coordonată, corespondenţă, cornet, costum, crea, credit, cristal, critică, cultiva, cultură, cupon, curbă, cuier, debit, debita, declinaţie, *degazifica,* degrada, *degradare,* demonstrativ, depozit, deriva, derivă, descrie, desena, *destrăma,* detaşa, detenţiune, determina, deveni, dezarticulat, diafragmă, dialect, dialectică, diamant, dictatură, difuzor, disciplină, discorda, discuta, dispozitiv, dispoziţie, disputa, distincţie, distinge, diversiune, dom, domina, dota, dublu, dur, efect, egalitate, elementar, emigraţie, emisferă, epocă, epuiza, eră, erou, esenţă, eter, etichetă, exerciţiu, expansiune, expediţie, explica, exploata, exploatare, expresie, expropria, extern, extremă, faţetă, favorit, fecund, fibră, filieră, filozof, fin, fineţe, fizic,



Scanned with OKEN Scanner

formal, fosfata, front, furnitură, gradaţie, gravitate, grilă, grup, grupa, identitate, ilustra, imprima, impune, *incarnat,* inerţie, infect, injector, inspira, inspiraţie, instala, institut, instrucţie, instrui, instrument, instrumentaţie, *interesat,* intern, interpret, intimitate, introduce, *introducere,* invocaţie, izola, încadra, *înclinat,* înregistra, jurisprudenţă, justifica, labirint, lamina, laminor, lansa, lanternă, legendă, legitima, lentilă, liberal, lichid, lichida, limb, limită, logică, lot, magazin, magnetism, manevră, manşetă, marca, marchiză, marş, marşrut, mecanism, meditaţie, membrană, memorie, minută, misiune, mister, mizerabil, mobil, mod, modalitate, model, modula, monta, montaj, mutaţie, naţionalitate, navetă, negativ, nervos, nişă, nobil, normal, normă, noţiune, obligaţie, obscur, obscuritate, *ocupat,* oficial, olimpic, ondula, *ondulat,* onor, onorific, orbită, ordonanţă, original, origine, paletă, palier, panoramă, panou, parchet, parfum, parolă, parter, pasaj, pasarelă, patrie, patron, pedală, pedagog, pegas, peliculă, pepinieră, percuţie, perlă, perspectivă, pigment, pilon, pilot, pionier, pirotehnie, plafon, planşă, planta, plasa, plasament, plastron, platou, plintă, poanson, pol, polar, politică, ponton, poseda, posesiune, postum, poza, preludiu, prepara, preparator, preparaţie, prescripţie, pretinde, prezumţie, primitiv, proces, procurator, profila, program, proiect, proporţie, proscrie, prostitua, protecţie, protocol, provoca, public, punctiform, rachetă, radiator, radier, radio, ramă, raţional, raţionaliza, reactor, reacţie, realiza, recoltă, recomanda, reconstitui, rectifica, reflecta, *reformat,* regenera, regie, relativ, relaţie, releu,



Scanned with OKEN Scanner

relief, repertoriu, repriză, retenţie, rol, rond, rotaţie, roză, rubrică, ruină, ruletă, rulou, sală, salon; scarifica, scarificator, scenă, schelet, secundă, *segregare, selecta,* semafor, semnal, *semnalizare,* senat, sensibil, sensibilitate, serpentină, siluetă, simpatic, simpatie, sincopă, sincronic, sindicat, sirenă, sistem, situaţie, social, soclu, solda, solicita, solidar, soluţie, sondaj, specula, spiră, spiritual, stand, static, staţie, steril, strangula, studio, *subsol,* substanţă, subtil, succeda, suită, sulfata, supapă, superb, *suprastructură,* suprima, surprinde, susceptibil, susceptibilitate, suspenda, sustrage, şarjă, tampon, tandem, temă, timpan, tip, titru, toleranţă, trahee, transmisiune, transpira, transport, transporta, trapez, tribunal, triunghi, trompă, tropic, tunică, tur, urnă, vampir, variometru, vedetă, vehicul, ventuză, vibra, viola, violent, violenţă, vira, virgulă, vizita, viziune, vulgariza, zero . . . . . 3 sensuri

Analogul tabelului 16.2. este, în acest caz, tabelul 16.4.

*Tabelul 16.4.*

| $s \geq 1$ | 1 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Nr. cuv. | 18 250 | 697 | 221 | 93 | 42 | 19 |
| ln (Nr. cuv.) | 9.81 | 6,55 | 5,40 | 4,53 | 3,74 | 2,94 |

Dacă nu se iau în considerare decît cuvintele-bază[1], primele 3 coloane ale tabelului 16.4. se modifică astfel:

*Tabelul 16.5.*

| $s \geq 1$ | 1 | 3 | 4 |
|---|---|---|---|
| Nr cuv.-bază | 15088 | 675 | 218 |
| ln (Nr. cuv.-bază) | 9,63 | 6,52 | 5,38 |

[1] În lista cuvintelor de origine „franceză", derivatele au fost evidenţiate prin caractere cursive.



Scanned with OKEN Scanner

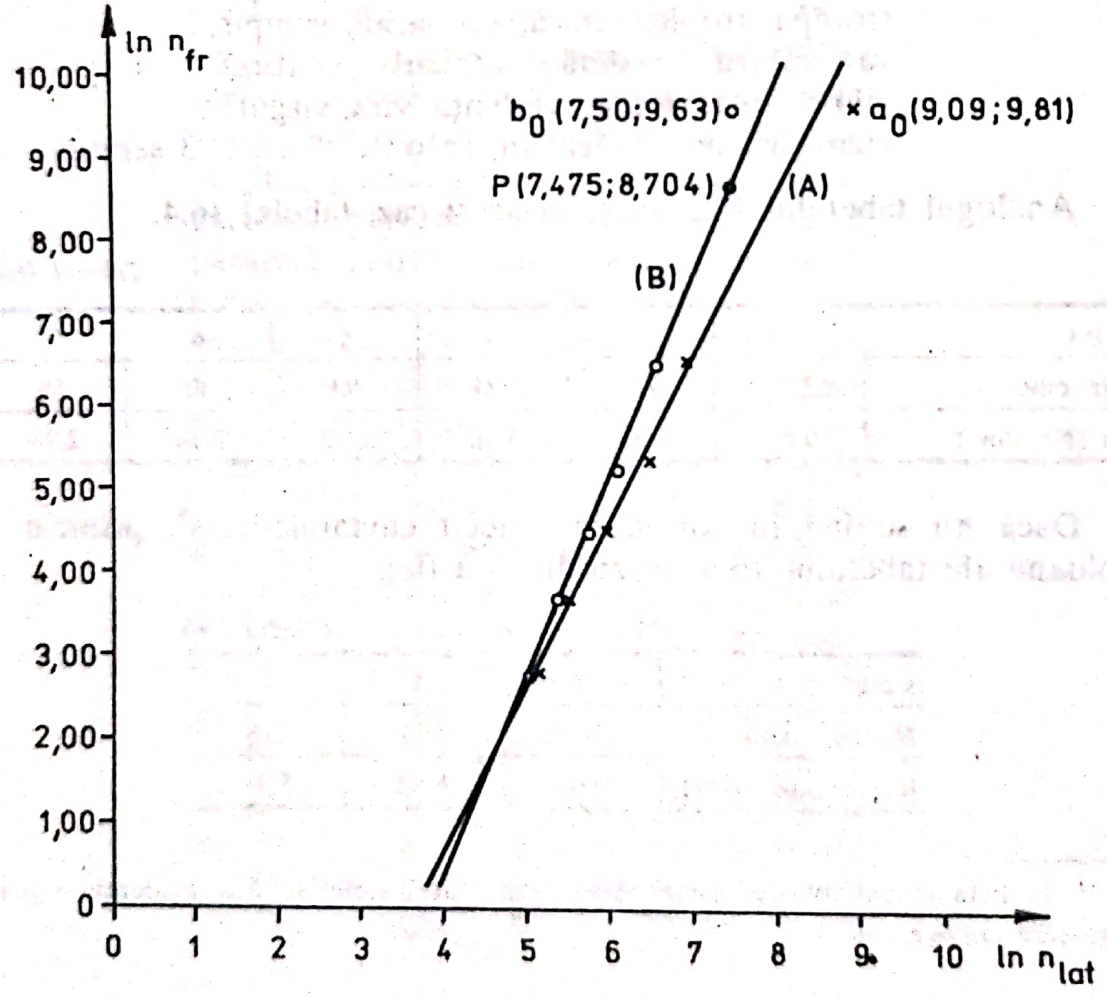
Corespondenţa dintre elementul latin şi cel „francez“ pentru valori identice ale numărului de sensuri $s$ este cea din tabelul 16.6.

*Tabelul 16.6.*

| $s \geq$ | | 1 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| TOTAL CUVINTE | $\ln n_{lat}$ | 9,09 | 6,99 | 6,49 | 5,99 | 5,54 | 5,18 |
| | $\ln n_{fr}$ | 9,81 | 6,55 | 5,40 | 4,53 | 3,74 | 2,94 |
| CUVINTE BAZĂ | $\ln n_{lat}$ | 7,50 | 6,62 | 6,14 | 5,78 | 5,40 | 5,08 |
| | $\ln n_{fr}$ | 9,63 | 6,52 | 5,38 | 4,53 | 3,74 | 2,94 |

Reprezentate într-un sistem de referinţă cartezian, avînd logaritmii numerelor de cuvinte de origine latină şi „franceză“ în abscisă şi, respectiv, în ordonată, cele două şiruri de perechi de valori din tabelul 16.6. conduc la două mulţimi de puncte (fig. 16.2) dispuse fiecare

FIG. 16.2



aproape coliniar, cu excepţia punctelor corespunzătoare inegalităţii $s \geq 1$. Ştiut fiind faptul că aceste din urmă puncte reflectă raportul dintre elementul latin şi cel „francez" în întreg vocabularul considerat (deoarece orice cuvînt posedă măcar un înţeles, ceea ce se notează prin $s \geq 1$), observăm că, întocmai cum totalitatea unităţilor de cele două provenienţe nu satisfăcea ecuaţia productivităţii, ea nu o satisface nici pe cea a polisemiei. Dreapta (A) din fig. 16.1., care ia în considerare atît cuvintele-bază cît şi derivatele, are ecuaţia:

$$ln\, n_{fr} = 1{,}942\, ln\, n_{lat} - 7{,}094 \qquad (16.1)$$

iar dreapta (B), ce se referă exclusiv la cuvintele-bază:

$$ln\, n_{fr} = 2{,}361\, ln\, n_{lat} - 9{,}081 \qquad (16.2)$$

Figura ne arată cît de mult se abat de la dreptele (A) şi (B) punctele $a_0$ (9,09 ; 9,81) şi $b_0$ (7,50 ; 9,63), care corespund totalităţii cuvintelor latine şi „franceze" din dicţionar.

Din analiza productivităţii (paragraful 10) am reţinut însă că, potrivit ecuaţiilor dreptelor caracteristice, logaritmul numărului total al cuvintelor-bază de origine latină ar trebui să fie $ln\, n_{lat} = 7{,}475$, iar al celor de origine „franceză" $ln\, n_{fr} = 8{,}704$. Cu alte cuvinte, în raport cu legităţile productivităţii lexicale, într-un vocabular conţinînd $e^{7{,}475} = 1\,763$ de cuvinte-bază moştenite din latineşte nu ar trebui să existe decît $e^{8{,}704} = 6\,026$ de cuvinte-bază de provenienţă „franceză". Tot restul, pînă la 15 021, cîte cuvinte-bază „franceze" figurează în DLRM, constituie aluviuni recente ce pot fi socotite drept încă neasimilate de sistemul *lexicogen* românesc.

Reprezentarea pe graficul din figura 16.2. a punctului $P$ (7,475 ; 8,704), corespunzător echilibrului numeric dintre componentele latină şi „franceză" indicat de modelul productivităţii evidenţiază o „coincidenţă" al cărei interes se va dovedi major: punctul $P$ cade[1] pe dreapta (B) de ecuaţie (16.2.). Aceasta înseamnă că şi potrivit modelului polisemiei lexicale un vocabular cuprinzînd 1 763 de cuvinte-bază latine ar trebui să conţină tot în jur de 6 026 de cuvinte-bază de origine „franceză". Tot ceea ce depăşeşte acest număr reprezintă împrumuturi ce nu au reuşit încă să se încadreze sistemului *semantogen* românesc. În cazul că nu constituie o simplă întîmplare (bănuială pe care analiza în continuare a altor aporturi o va înlătura), concordanţa celor două rezultate are o importanţă deosebită.

[1] Abaterea calculată are o valoare neglijabilă: 1,6%.



Scanned with OKEN Scanner

Ştiam deja (ne-o relevase studiul productivităţii) că sistemul lexical nu este dispus să primească imigranţi în orice proporţie. Legităţile ce guvernează raporturile cantitative dintre componentele etimologice ale lexicului nu permit încadrarea instantanee în *sistemul* vocabularului a tuturor elementelor vocabularului. Prelungind analogia pe care o sugerează termenul „imigrant", am putea spune că dintre cuvintele rezidente în limba română numai o parte posedă la un moment dat cetăţenia română. În cazul împrumuturilor „franceze" circa 9 000 dintre cele 15 000 de cuvinte consemnate în DLRM nu au reuşit încă să obţină acest statut preferenţial. Perspectiva lor de viitor este fie expulzarea, fie modificarea legii de acordare a cetăţeniei, în aşa fel încît condiţiile de naturalizare să fie uşurate, fie, şi acest scenariu e în mod evident cel mai probabil, deopotrivă relegarea unor indezirabili şi o anumită relaxare legislativă care să permită accesul la cetăţenie al unei părţi mai însemnate a imigranţilor. Nu trebuie să se înţeleagă deci că unităţile lexicale aflate în exces în raport cu prescripţiile celor două grupe de legi (referitoare la productivitate şi, respectiv, la polisemie) reprezintă corpuri străine ce trebuie neapărat eliminate, ci numai că poziţia lor momentană în limbă este nesigură, iar viitorul lor incert. O parte dintre ele riscă să intre în desuetudine, cum s-a întîmplat cu multe elemente turceşti şi neogreceşti. Dintre cele 1097 de cuvinte de origine turcească din DLRM, un număr de 226 poartă menţiunea *înv.* Faptul că, potrivit ecuaţiei (10.9.), ar fi trebuit să nu întîlnim decît $e^{6,234} = 510$ cuvinte provenite din turceşte, arată că procesul îndepărtării din limba vie a excedentului lexical de această obîrşie este departe de a se fi încheiat. O afirmaţie asemănătoare se poate face cu privire la componenta neogrecească. Din totalul de 624 de împrumuturi din neogreacă, 124 au ajuns să aibă statut de arhaisme, dar se poate anticipa că în viitor alte vreo 200 de cuvinte le vor împărtăşi soarta, deoarece ecuaţia (10.8.) indică drept probabilă prezenţa în limbă a cel mult 300 de unităţi de această origine. Viitorul de care vorbim nu este însă deloc apropiat. Dintre componentele istorice ale limbii române, practic numai fondul primordial latin nu prezintă elemente în exces. Chiar prezenţa slavă în dicţionar pare întrucîtva excedentară în raport cu nivelul prescris de ecuaţia (10.3). Dacă se observă însă că dintre cele 1 110 cuvinte-bază de origine slavă veche cuprinse în DLRM un număr de 98 sînt învechite iar 46 nu se mai folosesc decît regional, rămîn doar 966 de unităţi vii şi general dacoromâne, adică exact cîte preconizează modelul. Componenta veche slavă a vocabularului şi-a eliminat aşadar întreg balastul lexical neasimilat, ceea ce componentele turcă şi neogreacă nu au izbutit încă.



Scanned with OKEN Scanner

Homeostaza sistemului lexical necesită, după cum se vede, timpi de reglaj de ordinul unei jumătăţi de mileniu.

Dar eliminarea treptată a „surplusului" de cuvinte nu constituie singura cale de echilibrare a sistemului. Modificarea ecuaţiilor ce îl definesc este o a doua, cu nimic mai puţin importantă decît cea dintîi. O asemenea evoluţie e nu numai posibilă, ci şi istoriceşte inevitabilă. Ajunge să reamintim că, aşa cum s-a arătat, pantele dreptelor caracteristice sînt dependente de timp, ele scad cu trecerea vremii, afectînd relaţiile reciproce dintre componentele etimologice ale vocabularului. Pentru valoarea medie a numărătorului din relaţia (12.6.):

$$t = \frac{90 \div 95}{m^2}$$

panta dreptei caracteristice a aportului lexical $i$ variază cu timpul potrivit legii:

$$m_i = \frac{9{,}62}{\sqrt{t_i}}$$

unde $t_i$ este vechimea medie a cuvintelor de origine $i$ în raport cu momentul considerat (în cazul nostru: anul 1950).

La rîndul ei, panta dreptei ce exprimă raporturile dintre componentele $i$ şi $j$ va fi dublu dependentă de timp, întrucît (v. relaţia 13.4) ea are forma:

$$m_{ij} = \frac{m_i}{m_j} = \sqrt{\frac{t_j}{t_i}} \qquad (16.3.)$$

După trecerea unui interval de timp $\Delta t$, parametrul $m_{ij}$ va căpăta valoarea:

$$m_{ij} = \sqrt{\frac{t_j + \Delta t}{t_i + \Delta t}} \qquad (16.4.)$$

Variaţia simplă indicată de relaţia (16.4) nu este însă valabilă decît în cazul influenţelor ce au încetat să se exercite, adică al relaţiilor cu o limbă ale cărei contacte directe cu româna s-au încheiat. Dacă însă, şi aceasta este situaţia elementului „francez", împrumuturile continuă, vîrstele medii $t_i$ şi $t_j$ se modifică potrivit unor legităţi particulare, dependente de ritmul de pătrundere în româneşte a cuvintelor noi de



Scanned with OKEN Scanner

provenienţele respective, iar panta $m_{ij}$ se vede obligată să le urmărească modificările.

Toate aceste lucruri ne erau deja cunoscute. Ceea ce a intervenit nou acum şi este de natură să confere o greutate mult mai mare consideraţiilor anterioare e constatarea unei concordanţe surprinzătoare între limitele de acceptabilitate ale sistemului lexical în cele două ipostaze ale sale: cea de producător de termeni noi şi cea de producător de înţelesuri noi. Din ambele puncte de vedere, numărul de neologisme „franceze" pe care sistemul le poate încorpora într-un moment istoric dat, fără a-şi încălca norma proprie de funcţionare, este acelaşi. Rămîne să vedem dacă lucrurile stau la fel şi în ceea ce priveşte celelalte influenţe, fie ele mai vechi sau mai noi.

Un aport neologic într-un fel mai important ca cel „francez" (am văzut că furnizează 3 dintre cele 6 cuvinte nemoştenite posesoare a cîte cel puţin 10 sensuri) este cel latin savant. Pînă la nivelul de polisemie 3 el prezintă imaginea dată de lista ce urmează:

| | | |
|---|---|---|
| 1. | linie . . . . . | 13 sensuri |
| 2—3. | punct, spirit . . . . . | 11 sensuri |
| 4. | coroană . . . . . | 8 sensuri |
| 5—7. | ordin, reproduce, titlu . . . . . | 6 sensuri |
| 8—21. | capsă, colegiu, graţie, istorie, obiect, oficiu, particular, principiu, real, regină, spaţiu, susţine, virgin . . . . . | 5 sensuri |
| 22—53. | caz, concepe, conduce, culoare, decan, dedica, descinde, deşert (subst.), elibera, expune, extract, fals, gen, geniu, himeră, literă, materie, minor, natural, ordine, produce, propriu, propune, *puncta,* reduce, stabili, studiu, subiect, ştiinţă, termen, traduce, uz . . . . . | 4 sensuri |
| 54—106. | act, acut, aură, catenă, cauză, cens, clar, cloacă, *compunere,* concurge, *conducător,* cult (subst.), *depunere, descompune, descompus, desperat,* emite, for, imperiu, impuls, *încorona,* mediu, nerv, oferi, opune, oră, orizont, peduncul, perfect, pictură, prefectură, pretor, prim, probă, *produs,* profund, pronunţa, purpură, raţiune, *redus,* rege, *regulat,* regulă, *restabili,* roman, *sărbătoare,* stabil, sumă, superior, tezaur, triumfa, victorie . . . . | 3 sensuri |



Scanned with OKEN Scanner

Raporturile cantitative dintre fondul latin moştenit şi elementul latin savant pentru ambele variante (1. unităţi de orice fel; 2. numai cuvinte-bază) sînt sintetizate în tabelul 16.7.

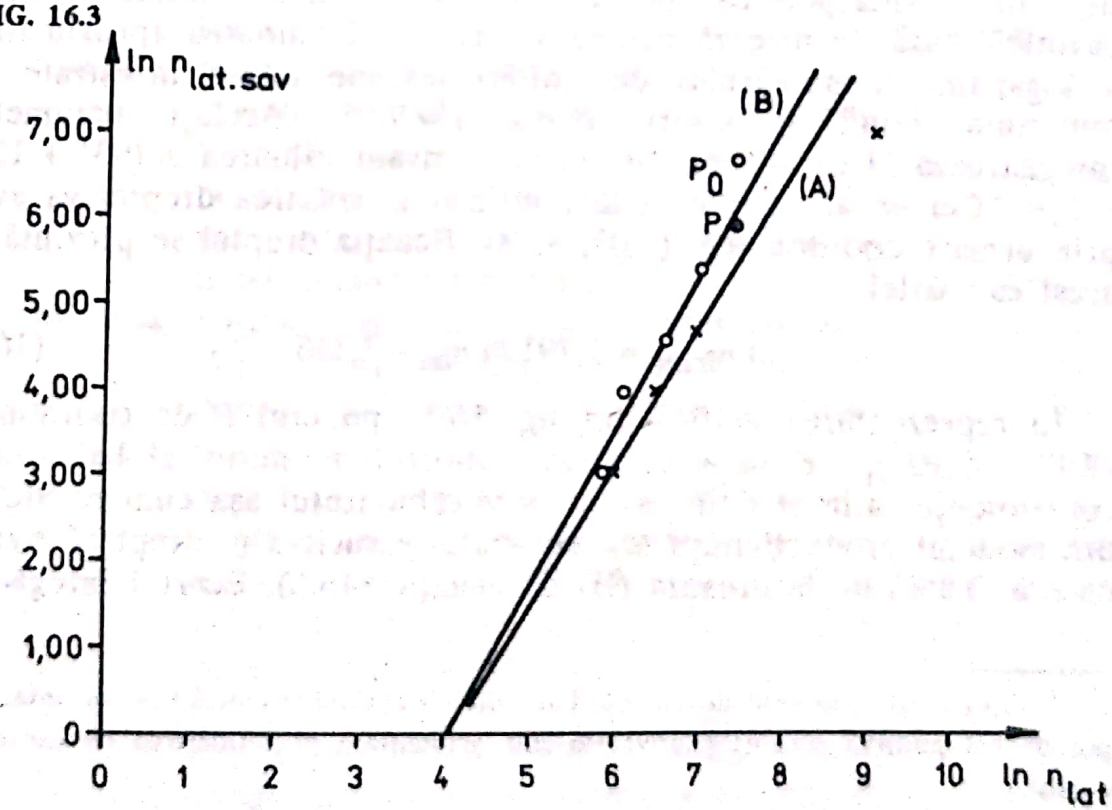
*Tabelul 16.7*

| $s \geq$ | | 1 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| TOTAL CUVINTE | $ln\ n_{lat}$ | 9,09 | 6,99 | 6,49 | 5,99 |
| | $ln\ n_{lat.sav}$ | 7,05 | 4,66 | 3,97 | 3,04 |
| CUVINTE-BAZĂ | $ln\ n_{lat}$ | 7,50 | 6,62 | 6,14 | 5,78 |
| | $ln\ n_{lat.sav}$ | 6,69 | 4,53 | 3,95 | 3,04 |

Dat fiind că în fig. 16.3. punctele corespunzătoare coloanei 1 sînt, ca şi în cazul elementului „francez", puternic deviante faţă de dreptele determinate de celelalte puncte (din motivul arătat al necoincidenţei dintre „rezidenţă" şi „cetăţenia" lexicală), trebuie să ne mulţumim numai cu cîte trei perechi de valori, ceea ce face ca trasarea dreptelor să fie destul de aproximativă. Pentru numărul total de cuvinte (termeni-matcă şi derivate), expresia analitică a dreptei ce descrie relaţia dintre elementul latin savant şi cea latin este:

$$ln\ n_{lat.sav} = 1{,}616\ ln\ n_{lat} - 6{,}598 \qquad (16.5)$$

FIG. 16.3






Scanned with OKEN Scanner

În ceea ce priveşte numai cuvintele-bază, vom încerca să îmbunătăţim precizia trasării dreptei adăugînd celor 3 puncte de care dispunem un alt patrulea, referitor la unităţile posesoare a cîte cel puţin 2 sensuri. Termenii-matcă de sursă latină savantă care au în româneşte două înţelesuri distincte sînt în număr de 125: acvatic, agramat, ascendent, ascensiune, atent, atrium, audia, aulă, boem, brut, cancelar, castel, catedră, catolic, cazual, centurie, circumferinţă, cisternă, claritate, climă, codice, combinaţie, comparaţie, compune, consiliu, conspect, consul, corupe, cretă, cronică, dac, dedica, delfin, demn, demnitate, depune, despera, dizolva, doctor, epistolă, exercita, extrage, falsitate, fascie, febră, fortifica, frecvenţă, fruct, german, germen, gramatică, gumă, gută, inamic, induce, insulă, lapidar, latin, laur, leonin, maur, medicină, medie, melisă, metal, minus, mizerie, multiplica, murmura, muză, muzică, nud, nutri, papalitate, pauză, peculiu, pendinte, plagă, plex, plus, pretoriu, preda, product, promite, promova, provincie, pur, regula, remite, repauza, retor, sacrifica, sacru, scripte, secol, serv, sigiliu, singular, solemnitate, specie, speluncă, speţă, stadial, stimul, sumar, tabelă, tabernacol, terminaţie, testament, topică, trac, trăda, triumf, triumfal, uman, umanitate, umil, unic, uniune, vacuum, varia, văl, violă, virtute, voce.

Pentru a nu mai relua examinarea întregului fond latin moştenit în căutarea cuvintelor cu două sensuri, ne vom folosi de graficul din fig. 16.1. Interpolarea pe curba (B), valabilă numai pentru cuvintele-bază, furnizează pentru abscisa $s = 2$ valoarea aproximativă a logaritmului numărului de unităţi lexicale latine înzestrate cu minimum două înţelesuri: $ln\, n_{lat}(2) = 7{,}05$. Acelaşi parametru înregistrează în cazul elementului latin savant valoarea $ln\ (93^1 + 125) = 5{,}38$. Cel de al patrulea punct utilizat la trasarea dreptei va avea prin urmare coordonatele (7,05; 5,38). Ecuaţia dreptei se prezintă în acest caz astfel:

$$ln\, n_{lt.sav} = 1{,}791\ ln\, n_{lat} - 7{,}246 \qquad (16.6)$$

În reprezentarea grafică din fig. 16.3., punctul $P$ de coordonate (7,475; 5,906), ce caracterizează raportul numeric global dintre constituenţii latin şi latin savant ai vocabularului aşa cum rezultă el din modelul productivităţii $M_p$, se abate întrucîtva (e drept că numai cu cca. 3,8%) de la dreapta (B) de ecuaţie (16.6). Ecartul înregistrat

---

[1] Acesta este numărul de cuvinte-bază din lista elementelor latine savante care posedă cel puţin 3 sensuri (derivatele sînt evidenţiate prin tipărirea cu caractere cursive).



Scanned with OKEN Scanner

nu pare suficient de concludent pentru infirmarea ipotezei formulate anterior, dar, firește, nu poate contribui nici la consolidarea ei. Frontiera vagă dintre neologismele latine savante şi cele „franceze" ar putea fi foarte bine responsabilă de această uşoară nepotrivire. Oricum, $P$ este mult mai apropiat de dreaptă decît punctul $P_0$ care corespunde numărului total efectiv de cuvinte-bază latine şi, respectiv, latine savante din DLRM.

Ţinînd seamă de tendinţa eliminării treptate din vocabular a prisosului lexical (în sensul discutat mai sus), ne aşteptăm ca în cazul elementului slav punctele $P$ şi $P_0$ să fie aproape suprapuse. Într-adevăr ele diferă extrem de puţin şi chiar ajung să coincidă în ipoteza renunţării la arhaisme şi regionalisme. Cum stau aceste puncte faţă de dreapta ce exprimă variaţia componentei vechi slave a vocabularului în raport cu fondul latin moştenit?

Pînă la nivelul de polisemie 3 inclusiv, cuvintele de obîrşie slavă veche prezintă următoarea ierarhie:

| | | |
|---|---|---|
| 1—2. | sfînt, val . . . . . | 9 sensuri |
| 3—6. | coş, lovi, *potrivi,* prost. . . . . | 8 sensuri. |
| 7—15. | cinste, *coborî,* cumpăni, pod, pomeni, prag, *răsuci,* strajă, treabă. . . . . | 7 sensuri |
| 16—29. | *atîrna,* coajă, *deosebit,* gol, *împleti,* jghiab, ocol, rînd, tablă, topi, trăi, trezi, trup, ţintă . . . . . | 6 sensuri |
| 30—61. | babă, brazdă, canon, citi, cocoş, croi, *dezveli,* dovedi, drojdie, duh, izvor, *lovitură,* muncă, opri, *opritoare,* păzi, poală, pofti, *potrivit,* primi, răzbi, risipi. robi, sită, slab, *slăbiciune, sucit,* tabără, taină, tocmi, ţigan, vreme. . . . . | 5 sensuri |
| 62—144. | *bogăţie, călit,* cămin, ceas, *cernit,* ciolan, cîrpi, clăti, clopot, cumpănă, doborî, dungă, fală, glas, goni, gunoi, hîrtie, ispită, *ispiti,* izbi, *împletit,* întîmpina, învîrti, lipi, mărturisi, meni, milă, munci, năpastă, năvăli, nevoie, obraz, *obrăzar,* omorî, *otrăvit,* oţel, păianjen, pinten, pîclă, plesni, *plesnitoare,* plug, ponos, popă, porni, *pornire,* povară, *prăpăstios, primire, privire,* rac, rîndui, rob, ropot, scîrbă, scoică, *scumpete,* sfinţi, slavă, slugă, sminti, *snopi,* soroc, *spori, sprijini,* | |



Scanned with OKEN Scanner

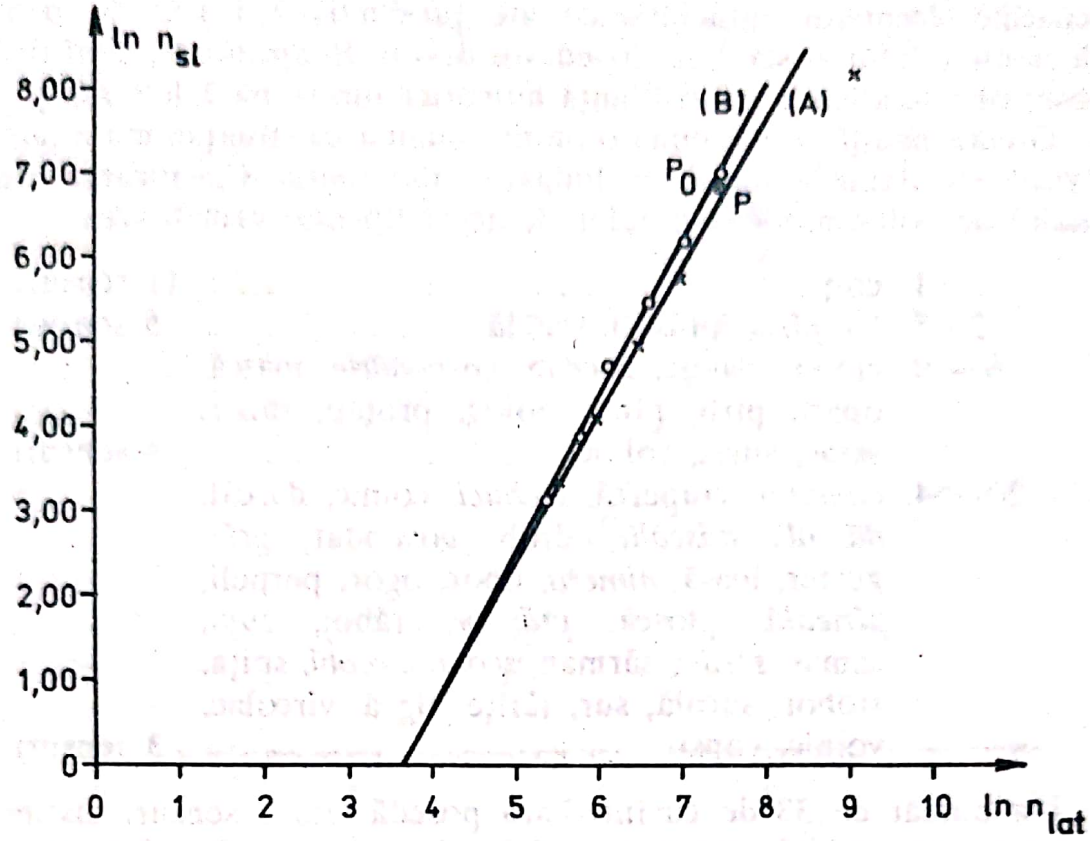
*stăvilar,* strună, suci, tainic, tălmăci, toiag, trăsni, trîmbă, ţeapăn, veac, verigă, vesel, vîlvă, vîrf, vîrstă, voie, *zări,* zmeu. . . . . 4 sensuri

145—353. apostol, arhiepiscopie, *aţinti,* boală, bob, bogat, bold, breaslă, *breslaş,* candelă, caznă, căli, ceată, cerneală, cerni, cinsti, ciocan, *ciocănar, ciocăni, cititor,* cîrn, *clătina,* clocoti, *clocotitor,* cobe, *cocoşel, cojit,* colac, *colăci,* colind, *cositură, crăiţă, croitor,* dărui, deosebi, *dezrobi,* dezvălui, dihanie, dîrz, *dobîndă,* dobîndi, drag, dragoste, *drăguţ,* drum, *fălos,* gît, greşi, *grindei, gunoi,* har, iaz, iute, *iuţeală, iuţi,* izbîndi, izvod, izvodi, *încleia, încleşta, încropi, întîmpinare, înţepeni, învălui, învrednici, îngrădi,* jelanie, jeli, jertfă, leit, *lipici, lipicios, lopătar, miros, mirosi,* mîndru, *mînjeală,* mîzgă, *mlădia, mlădios,* movilă, năduf, năpădi, năpraznic, năpusti, *năvodar,* nemernic, noroc, obicei, obidă, obşte, *ocoli,* ogradă, *orînduială, orînduire,* osebi, osteni, otrăvi, *oţeli,* ovăz, patriarh, pecete, pitac, pîndi, *pîrgă,* plată, plăti, *plăviţ, pleasnă,* pleşuv, *poci, poftire,* poftori, pogan, poiată, pojar, pomană, *pomenire, poprit, poruncă,* porunci, potop, potrivnic, poveste, praf, prapur, praştie, prăji, *prevesti, priveliţte,* privi, *prundar,* puzderie, răs-, răscol, răscoli, *răscolit,* răspîndi, *rînduială, robie,* rod, *scîrbi, scîrbit, scîrbos, sclipet,* scorpie, scump, sfat, *sfătui,* sfîrşi, *sfîrşit,* silă, sili, *slăbi,* slujbă, sluji, *solie,* spoi, spor, sprijin, staroste, stihie, *stîrcit,* straşnic, *struji,* struni, ştirb, *ştirbi,* tăbărî, tăbliţă, teasc, temei, tîngui, tîrg, tîrî, toci, tocmai, *tocmeală, trai,* treaz, trebui, trosni, *ţigănie, ţigănuş, ţinti,* ţintui, vedenie, *vesti,* vină, vinovat, *vîrfui,* vîrî, voievod, *volnici,* vornic, *vornicel,* vrăji, *vrednicie,* zare, zăbavă, zăbovi, zăbrea, zămisli, zăvor, zdravăn, zidire, zidit, zîmbi, zori . . . . . 3 sensuri



Scanned with OKEN Scanner

FIG. 16.4




Adăugînd, ca şi în cazul aportului latin savant, un punct obţinut prin interpolare, corespondenţa slav-latin are forma din tabelul 16.8.

Dreapta (A), referitoare la totalitatea cuvintelor, trasată prin punctele 3, 4, 5 şi 6 (fig. 16.4) are ecuaţia:

$$ln\ n_{sl} = 1{,}753\ ln\ n_{lat} - 6{,}382 \qquad (16.7)$$

iar dreapta (B), care priveşte numai cuvintele-bază, dusă prin punctele 2′, 3′, 4′, 5′ şi 6′ — ecuaţia:

$$ln\ n_{sl} = 1{,}860\ ln\ n_{lat} - 6{,}832 \qquad (16.8)$$

*Tabelul 16.8*

| $s \geq$ | | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| TOTAL CUVINTE | $ln\ n_{lat}$ | 9,09 | — | 6,99 | 6,49 | 5,99 | 5,54 |
| | $ln\ n_{slav}$ | 8,24 | — | 5,87 | 4,97 | 4,11 | 3,37 |
| CUVINTE BAZĂ | $ln\ n_{lat}$ | 7,50 | 7,05 | 6,62 | 6,14 | 5,78 | 5,40 |
| | $ln\ n_{slav}$ | 7,05 | 6,20 | 5,48 | 4,73 | 3,89 | 3,14 |



Scanned with OKEN Scanner

Faţă de această ultimă dreaptă, punctul $P$ (7,475 ; 6,878) dedus din ecuaţiile dreptelor caracteristice ale productivităţii lexicale pentru elementele latin şi slav[1] se dovedeşte destul de apropiat[2], confirmînd observaţia relativă la concordanţa numerică dintre modelele $M_p$ şi $M_s$.

O coincidenţă şi mai bună o oferă examinarea situaţiei materialului lexical cu dublă etimologie bulgaro-sîrbocroată. Clasificarea după gradul de polisemie a cuvintelor de acest tip este următoarea:

| | | |
|---|---|---|
| 1. | colţ . . . . . | 11 sensuri |
| 2—5. | ba, *pîrlit,* prăvăli, schelă. . . . . | 5 sensuri |
| 6—19. | ciopli, clăbuc, cocean, *gospodărie,* matcă, opări, pîrli, pluti, polog, proţap, răzor, sărac, surlă, voinic . . . . . | 4 sensuri |
| 20—54. | *cioplitor,* ciupercă, *clăbuci,* colnic, dascăl, *dăscăli, dăscălie,* drob, gospodar, grăi, guşter, leasă, *nimerit,* obor, ogor, perpeli, *pîrlealâ,* ploscă, *prăşitor,* răboj, rujă, samar, *sărăci,* sărman, scoabă, *scobi,* spiţă, stobor, sucală, sur, tărîţe, tigvă, vîrcolac, voinic, vopsi . . . . . | 3 sensuri |

Un număr de 33 de cuvinte-bază posedă cîte 2 sensuri distincte. Acestea sînt: babiţă, busuioc, cetină, cîrpă, cladă, desetină, dogori, gaidă, gîrbiţă, grădină, maică, momi, opăci, pleaşcă, potcoavă, pravăţ, prăsi, prăşi, puf, răsplăti, scrîşni, solniţă, studeniţă, tigaie, tolbă, topilă, tropot, tuli, ţesală, zădări, zăduf, zlătar, zob.

*Tabelul 16.9*

| $s \geq$ | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| TOTAL CUVINTE | $ln\ n_{lat}$ | 9,09 | — | 6,99 | 6,49 |
| | $ln\ n_{bg.scr}$ | 6,37 | — | 3,99 | 2,94 |
| CUVINTE BAZĂ | $ln_{lat}$ | 7,50 | 7,05 | 6,62 | 6,14 |
| | $ln\ n_{bg.scr}$ | 5,15 | 4,33 | 3,76 | 2,83 |

Materialul avut la dispoziţie este, după cum se vede, cam sărac, fapt ce nu poate rămîne fără urmări asupra siguranţei rezultatelor. Dacă în cazul elementului „francez" sau al celui vechi slav dispuneam,

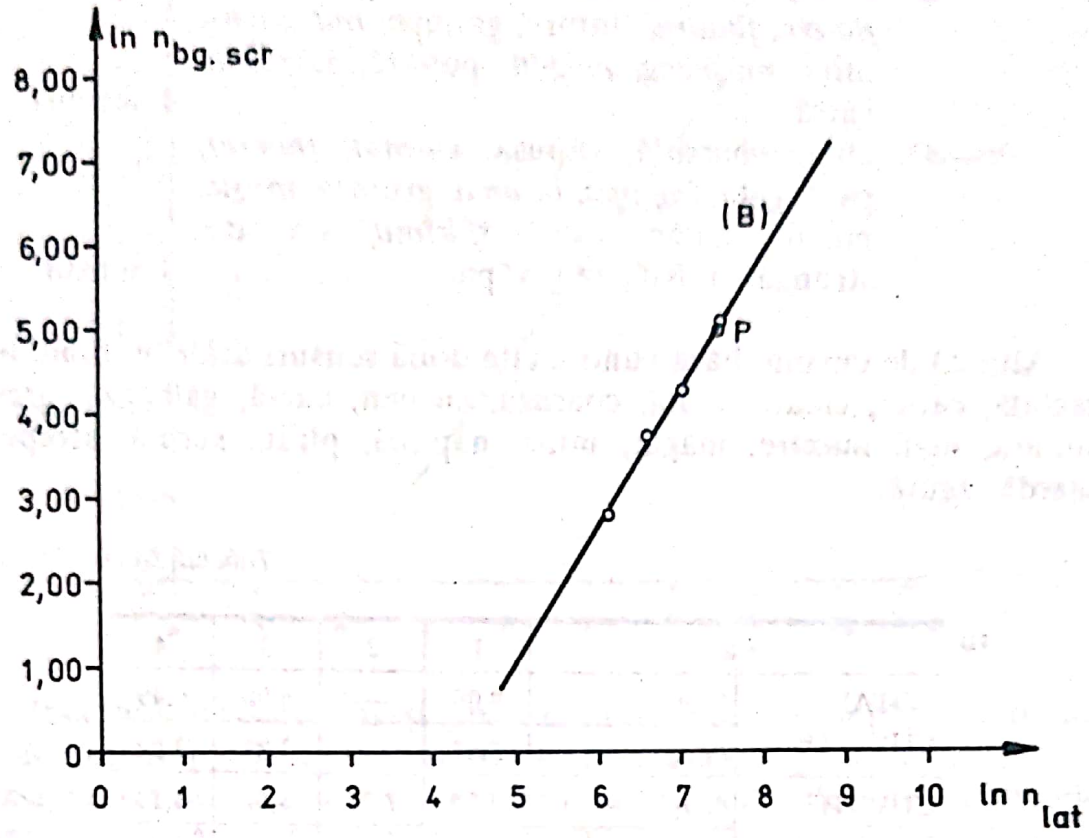
[1] Coordonatele acestui punct sînt termenii liberi ai ecuaţiilor dreptelor caracteristice: $ln\ n_{lat} = 7{,}475 - 0{,}230\ d$ şi $ln\ n_{sl} = 6{,}878 - 0{,}312\ d$.

[2] Ecartul său faţă de valoarea calculată este de 2,7%.



Scanned with OKEN Scanner

FIG. 16.5.




de bine de rău, de 4—5 puncte, aşezate astfel încît caracterul liniar al corelaţiei reieşea destul de limpede, de data aceasta liniaritatea nu mai este deloc evidentă. Ba chiar, pentru situaţia totalităţii cuvintelor (termeni-bază şi derivate), ea nu mai poate fi susţinută cu date obiective, întrucît cunoaşterea a numai două puncte (3 şi 4) este total insuficientă pentru deducerea alurii curbei căreia acestea îi aparţin.

În ceea ce priveşte cuvintele-bază, amplasarea celor 3 puncte de care dispunem (fig. 16.5) este compatibilă cu o relaţie de forma:

$$ln\ n_{bg.scr} = 1{,}620\ ln\ n_{lat} - 7{,}057 \qquad (16.9)$$

Faptul că punctul $P(7{,}475 : 5{,}051)$ cade pe această dreaptă (cu o abatere de numai 0,03%!) e de natură să ne sporească încrederea în rezultatul de mai sus.

Asemănătoare din acest punct de vedere este situaţia substratului. După numărul lor de înţelesuri, cuvintele „comune cu albaneza" se clasifică astfel:

1. de (conj.) . . . . . 13 sensuri
2. moş . . . . . 9 sensuri
3. ţarc . . . . . 7 sensuri
4—8. brîu, curma, fluier, *găti*, jumătate. . . . . 5 sensuri



Scanned with OKEN Scanner

9—25. *aburi,* bucura, buză, cioc, copil, *copilă, fluiera, flutura,* fluture, ghimpe, *măzăriche,* mîţă, *moşneag, pregăti,* pupăză, *sugruma,* vatră . . . . . 4 sensuri

26—45. abur. *abureală,* căpuşă, *curmat, fluierar,* gard, gata, *îngropa, înţărca,* grumaz, *moşie,* mugur, pălărie, raţă *sfărîma,* sîmbure, strungă, traistă, ţap, văpaie. . . . . 3 sensuri

Alte 23 de cuvinte-bază cunosc cîte două sensuri diferite: brad, buc, căciulă, cătun, cioară, ciută, coacăză, curpen, cursă, gălbează, gresie, gudura, mal, mazăre, măgar, mînz, năpîrcă, pîrău, scrum, strepede, zgardă, zgură.

*Tabelul 16.10*

| $s \geq$ | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| TOTAL CUVINTE | $ln\ n_{lat}$ | 9,09 | — | 6,99 | 6,49 |
| | $ln\ n_a$ | 6,05 | — | 3,81 | 3,22 |
| CUVINTE BAZĂ | $ln\ n_{lat}$ | 7,50 | 7,05 | 6,62 | 6,14 |
| | $ln\ n_a$ | 4,44 | 3,95 | 3,37 | 2,77 |

Din nou pentru ansamblul cuvintelor nu dispunem de suficiente elemente pentru a construi o ecuaţie credibilă (curba dusă prin numai două puncte poate fi de orice grad şi a o considera apriori drept o dreaptă constituie un abuz). Cele trei puncte care descriu situaţia relativă a cuvintelor-bază latine şi autohtone sînt însă dispuse coliniar pe dreapta de ecuaţie:

$$ln\ n_a = 1{,}282\ ln\ n_{lat} - 5{,}102 \qquad (16.10)$$

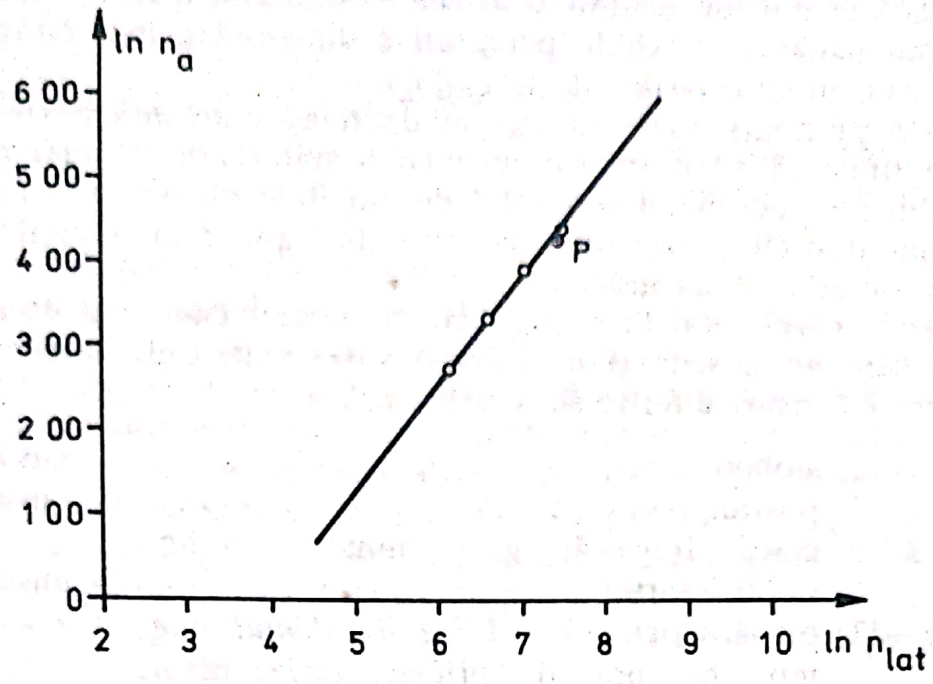
Punctul $P(7{,}475\ ;\ 4{,}329)$ indicat de ecuaţiile dreptelor caracteristice (10.1) şi (10,2) ca punct de echilibru teoretic global dintre cele două aporturi se află şi de data aceasta aproape de dreaptă (fig. 16.6.).

În legătură cu cele de mai sus şi, cu atît mai mult, cu cele ce vor mai urma în cuprinsul acestui capitol, este cazul să recunoaştem că cititorul, pe care ni-l dorim cît mai exigent, are toate motivele să fie nemulţumit de precizie redusă a determinărilor noastre. Comparativ cu acurateţea deosebită a modelului productivităţii $M_p$, exactiatea şi siguranţa modelului polisemiei $M_s$ pe care încercăm să îl construim



Scanned with OKEN Scanner

FIG. 16.6.




aici lasă mult de dorit. Însăşi „ipoteza liniarităţii", pe deplin confirmată în domeniul producţiei de derivate, prezintă un grad de certitudine considerabil mai scăzut cînd se referă la producţia de sensuri lexicale. Vinovată de această situaţie este penuria relativă de date de bază, împotriva căreia nu putem lupta, chiar dacă unele artificii de calcul ne dau iluzia îmbogăţirii fondului informaţional iniţal. Astfel, confirmarea liniarităţii relaţiei dintre productivitatea aporturilor latin şi vechi slav se baza pe existenţa unei mulţimi de 15 puncte aproape coliniare: în cazul indicatorului polisemie, nu dispunem în acelaşi scop decît de 4 (eventual 5, dacă se ţine seamă şi de punctul obţinut prin interpolare) puncte distincte; dreapta ce caracterizează raporturile dintre fondul latin şi cel autohton era construită în primul caz cu ajutorul a 9 puncte, în cel de al doilea cu numai 3 şi aşa mai departe. Incertitudinea care umbreşte rezultatele astfel obţinute va deveni încă şi mai supărătoare cînd vom trece la analiza unor componente etimologice ale vocabularului numericeşte inferioare celor de mai sus. Acceptînd însă în mod deliberat aceste limite inerente, preferăm ca, în ciuda impreciziei relaţiilor obţinute, să nu abandonăm obiectivul elaborării modelului $M_S$, deoarece socotim de preferat o cunoaştere parţială, amendabilă în perspectivă, a unor realităţi, necunoaşterii lor absolute. Măcar şi numai din punct de vedere strict metodologic, construirea unui model, fie el cît de imperfect pentru moment, al structurii polisemice a vocabularului îşi are importanţa sa, iar posibilitatea comparării acestuia cu modelul productivităţii lexicale,



Scanned with OKEN Scanner

chiar dacă nivelul de acurateţe al celor două constructe matematice este neasemănător, deschide perspective dintre cele mai atrăgătoare pentru viitorul cercetărilor de lexicologie.

Aceste precizări odată făcute, în deplina cunoştinţă a riscurilor, vom continua să adăugăm cărămizi cu o stabilitate tot mai precară edificiului $M_S$, sacrificînd scrupulul de exacitate obiectivului prioritar al completitudinii sistemului, condiţie sine qua non a confruntării structurilor celor două modele.

Ca şi în cazul modelului $M_P$, tratăm separat cuvintele de origine bulgară care nu posedă şi un etimon sîrbocroat. Cele înzestrate cu minimum 3 sensuri diferite se ierarhizează astfel:

1. slobod . . . . . 7 sensuri
2. pustiu. . . . . 6 sensuri
3—9. crac, drăgaică, griji, nimeri, odihni, *slobozi,* vîrtej . . . . . 5 sensuri
10—23. blană, cloci, *crăcană, îngriji, înlănţui,* lanţ, miji, na, prăpădi, pricină, raclă, răzor, sărac, sloi . . . . . 4 sensuri
24—55. borceag, clacă, comarnic, copcă, desluşi, *desluşit,* drob, grozav, *hărnici,* horă, icni, izbucni, *îngrijire, îngrijit,* jeg, meliţa, nadă, *nălucă,* ocniţă, *odihnă, odihnit,* păstra, piperniţă, podmol, prilej, *pustii,* scrob, slăbănog, stărui, vîrcolac, vopsi, zălud . 3 sensuri

Cuvintele-bază care au cîte două înţelesuri sînt în număr de 77: arap, baraboi, branişte, breaz, brei, bujor, canara, caracudă, călifar, căpiţă, cipic, ciudat, cîrmaci, clenci, clisă, clonţ, cloşcă, cosiţă, covergă, creangă, cremene, cută, dîrdoră, dîrjală, dumeri, galeş, gîdila, gîlcă, gligan, greblă, grijă, hapcă, hîrzob, iată, jumară, jupui, lele, magherniţă, mezdri, nisip, nojiţă, obială, oişte, opincă, ostreţ, oţări, papară, păstrăv, pelin, pită, plivi, pogon, pripor, prîsnel, proclet, pustie, rariţă, rînji, rîşniţă, rovină, rudă, sabie, sanie, scripete, sfoiag, sfredel, suman, suveică, tîrsînă, tor, torişte, ulei, vîlnic, vîslă, vrană, zăpor, zglăvoc.

Tabelul 16.11 priveşte, exclusiv, cuvintele-bază, dat fiind că numai pentru acestea dispunem de trei puncte care să ateste, pe cît este posibil, liniaritatea relaţiei.

*Tabelul 16.11*

| $s \geq$ | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| CUVINTE BAZĂ | $ln\ n_{lat}$ | 7,50 | 7,05 | 6,62 | 6,14 |
| | $ln\ n_{bg}$ | 5,66 | 4,80 | 3,81 | 3,04 |



Scanned with OKEN Scanner

Dreapta dusă prin punctele 2,3 şi 4 (fig. 16.7) are ecuaţia:

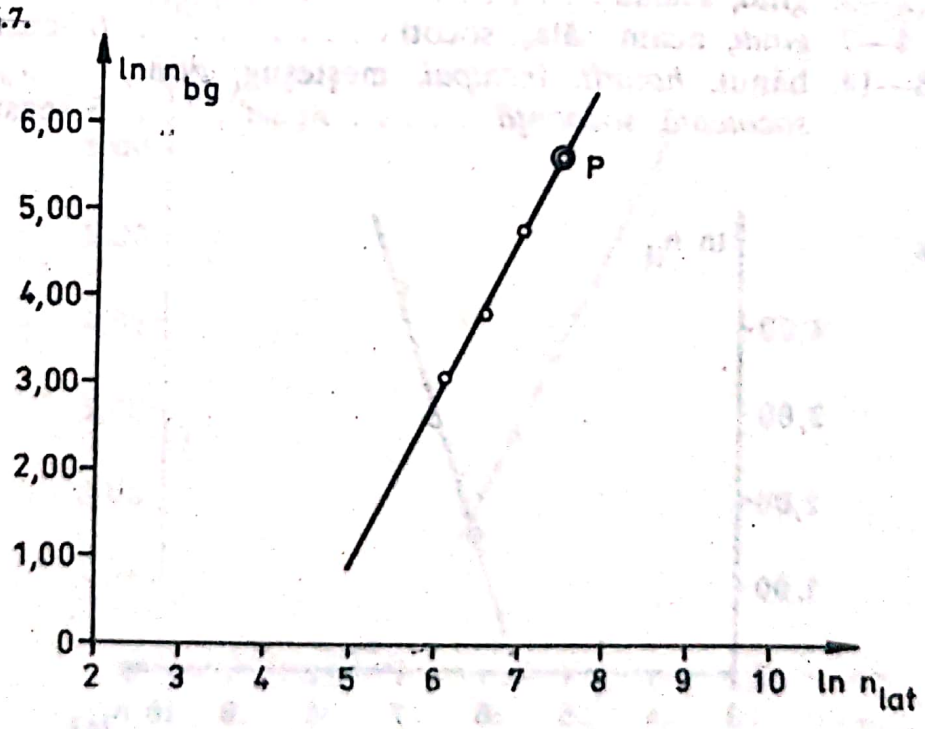
$$ln\ n_{bg} = 1{,}901\ ln\ n_{lat} - 8{,}669 \qquad (16.11)$$

Şi de această dată, punctul $P(7{,}475\,;\,5{,}597)$, furnizat de modelul productivităţii, cade pe dreaptă (cu o abatere de numai 1%).

Ultima situaţie în care mai dispunem încă de trei puncte pentru trasarea dreptei caracteristice priveşte raporturile dintre fondul latin şi împrumuturile din limba italiană. O nouă dovadă a faptului că neologismele, fie ele de provenienţă romanică sau latină savantă, sînt mult mai bine situate sub aspectul polisemiei decît sub cel al productivităţii. Vom vedea mai departe că aporturi lexicale superioare ca vechime şi productivitate celui italian, cum ar fi elementul ucrainian sau cel sîrbocroat, au furnizat atît de puţine unităţi polisemice încît va fi necesară recurgerea la un artificiu pentru a se putea obţine minimumul de 2 puncte necesare definirii unei drepte. Elementul italian mai permite încă trasarea dreptei cu ajutorul a trei puncte de coordonate (7,05 ; 4,39), (6,62 ; 2,89) şi (6,14 ; 1,61), corespunzătoare următoarei clasificări a cuvintelor-bază:

| | | |
|---|---|---|
| 1. | trata . . . . . | 5 sensuri |
| 2—5. | indice, respinge, siguranţă, tratament . . | 4 sensuri |
| 6—18. | bariton, bas, burghez, calendar, căpitan, cifră, codă, interes, medalion, papagal, piaţă, revistă, tapet . . . . . | 3 sensuri |

FIG. 16.7.






Scanned with OKEN Scanner

19—81. agenţie, alivanta, alto, amor, ancoră, armată, asalt, bagatelă, balsam, bancă, barbetă, bastard, baston, bilet, cameră, campion, campionat, canţonetă, casă 2, caustic, cavalerie 2, cifra, ciocolată, coloratură, costa, danez, exploda, fante, fascist, gelos, gelozie, impiegat, iredentist, italian, locotenent, maestru, manoperă, momentan, monopolist, operă (muz.), portofoliu, portughez, principe, raşchetă, raţiona, restanţă, reuşi, schiţa, schiţă, scont, sconta, solo, sorginte, stindard, studia, suplini, terţet, tirolez, tratat, trataţie, trio, vinci, violă, vulcanic . . . . 2 sensuri

Ecuaţia dreptei din figura 16.8 este:

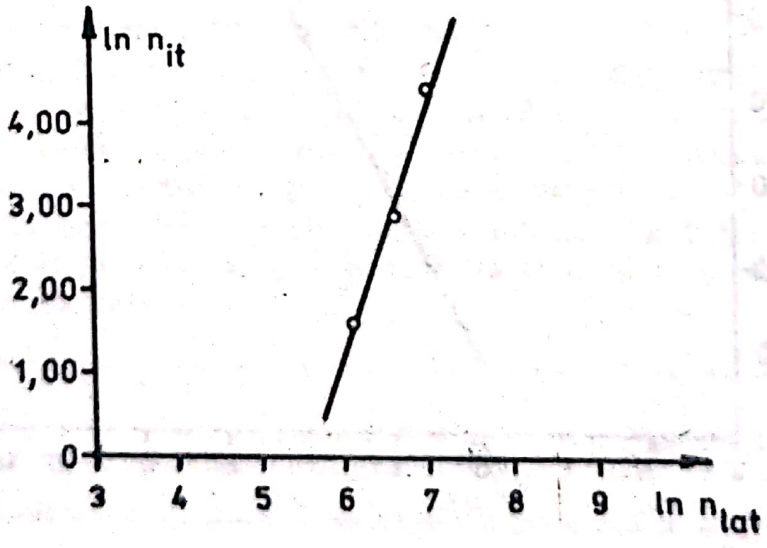
$$\ln n_{it} = 2.986 \ln n_{lat} - 16{,}754 \qquad (16.12)$$

Pentru aporturile maghiar, grecesc şi onomatopeic trebuie să ne mulţumim cu numai 2 puncte cunoscute.

Cuvintele de origine maghiară posesoare a cîte cel puţin 3 sensuri sînt:

1. chip . . . . . 9 sensuri
2—3. gînd, seamă . . . . . 7 sensuri
4—7. *gîndi,* neam, sălaş, socoti . . . . . 6 sensuri
8—14. bănui, *hotărît, închipui,* meşteşug, pont, *socoteală, socotinţă* . . . . . 5 sensuri

FIG. 16.8.



15—24. bănat, *cheltuială,* corlată, *hotărî, întîlni,* majă, pecetlui, tobă, ţarcă, zăbală . . . . 4 sensuri

25—53. ademeni, alcătui, alean, *buduroi,* chibzui, *chinui,* cotlon, dorobanţ, fel, gazdă, gingaş, *gîndire,* guler, hătui, hălădui, *hărţui,* joi (vb.), meşter, mistui, *nesocotit,* oraş, pandur, panglică, pîrcălab, pîrgar, prubului, şuvoi, vamă, viclean . . . . . 3 sensuri

*Tabelul 16.12*

| s ≥ | | 1 | 3 | 4 |
|---|---|---|---|---|
| CUVINTE-BAZĂ | $ln\ n_{lat}$ | 7,50 | 6,62 | 6,14 |
| | $ln\ n_{magh}$ | 6,11 | 3,69 | 2,77 |

Dreapta dusă prin punctele 3 şi 4 are ecuaţia:

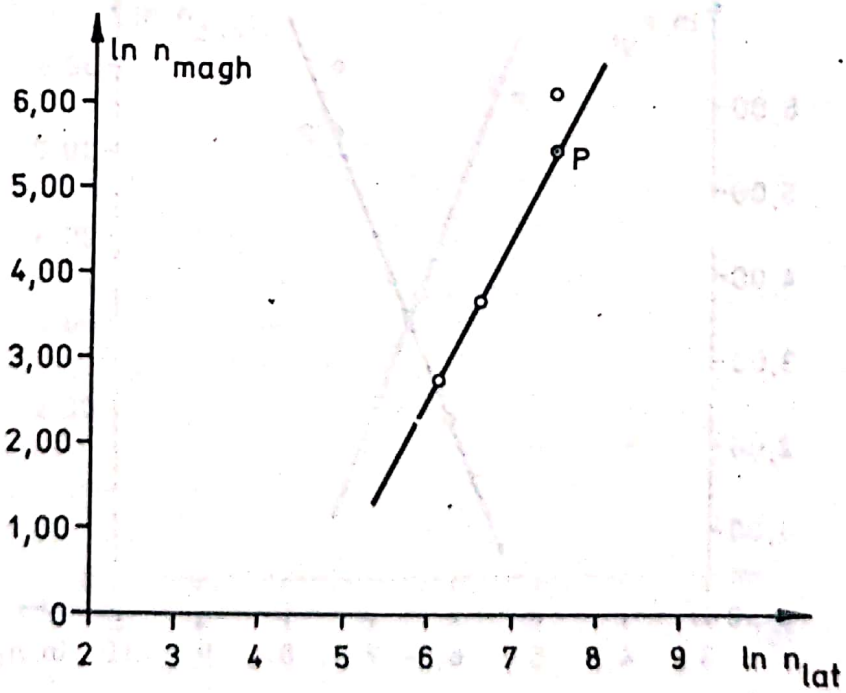
$$ln\ n_{magh} = 1{,}914\ ln\ n_{lat} - 8{,}981 \qquad (16.13)$$

În raport cu poziţia calculată cu ajutorul acestei expresii, ordonata punctului $P(7{,}475\ ;\ 5{,}457)$ se abate cu numai 2,4% (v. fig. 16.9).

Clasificarea termenilor de origine grecească, pînă la nivelul de polisemie 3, este:

1. pat . . . . . 7 sensuri
2. sigur . . . . . 6 sensuri

FIG. 16.9



3—5. chivernisi, *lipsă,* lipsi . . . . . 5 sensuri

6—12. mană, mărgăritar, partidă, piper, sîrmă, stemă, *zugrăveală* . . . . . 4 sensuri

13—41. *asigura, asigurare,* arhierie, bulin, condei, dană, *epitropie,* fasole, *folosi, garofiţă,* logofăt, măcar, noimă, oropsi, patriarhie, pălimar, *pătimaş, piperat,* plăsmui, portocală, proaspăt, saramură, sfoară, sosi, stol, tipar, ţaţă, zaharisi, zodie . . . . . 3 sensuri

*Tabelul 16.13*

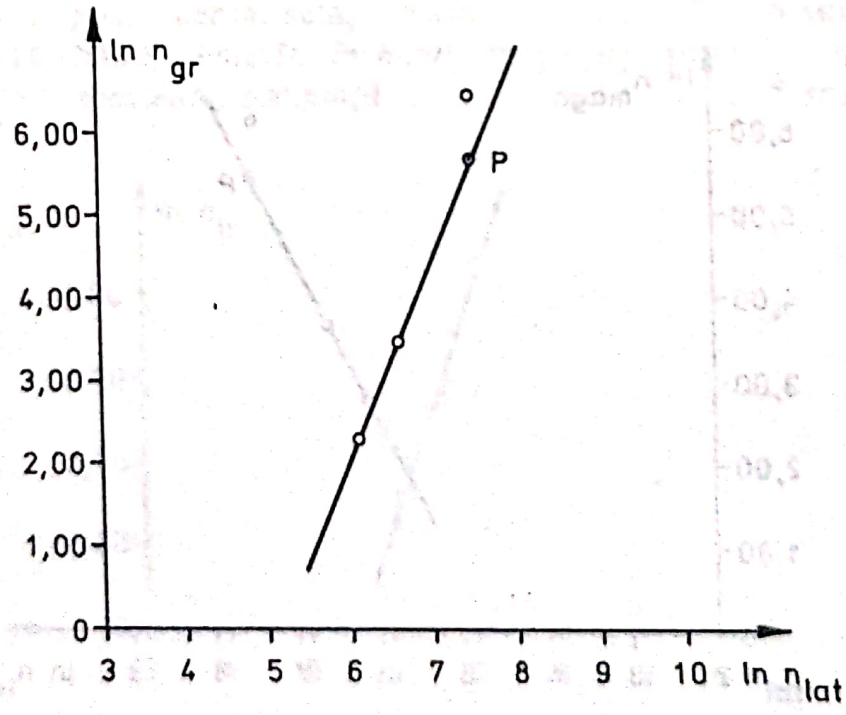
| $s \geq$ | | 1 | 3 | 4 |
|---|---|---|---|---|
| CUVINTE-BAZĂ | $ln\ n_{lat}$ | 7,50 | 6,62 | 6,14 |
| | $ln\ n_{gr}$ | 6,43 | 3,47 | 2,30 |

Trasată prin punctele 3 şi 4, dreapta din figura 16.10 are ecuaţia:

$$ln\ n_{gr} = 2{,}431\ ln\ n_{lat} - 12{,}625 \qquad (16.14)$$

Ecartul punctului $P(7{,}475\,;\,5{,}677)$ în raport cu poziţia calculată cu relaţia (16.14) este de 2,3%.

**FIG. 16.10**






Scanned with OKEN Scanner

Formaţiile lexicale onomatopeice se ierarhizează după numărul lor de înţelesuri astfel:

| | | |
|---|---|---|
| 1. | pica . . . . . | 8 sensuri |
| 2. | hop . . . . . | 7 sensuri |
| 3—4. | bufni, hîrîi . . . . . | 5 sensuri |
| 5—8. | ciovică, durui, ţîrîit, vuiet . . . . . | 4 sensuri |
| 9—28. | barangă, bîzîii, brr, cîrîi, clefăi, colcăi, ha, hai, hait, hei, hodorogi, plescăit, pocni, scîrţ, sfîrîi, toc, turui, ţîşni, ţîţîi, zbîrnîi | 3 sensuri |

*Tabelul 16.14*

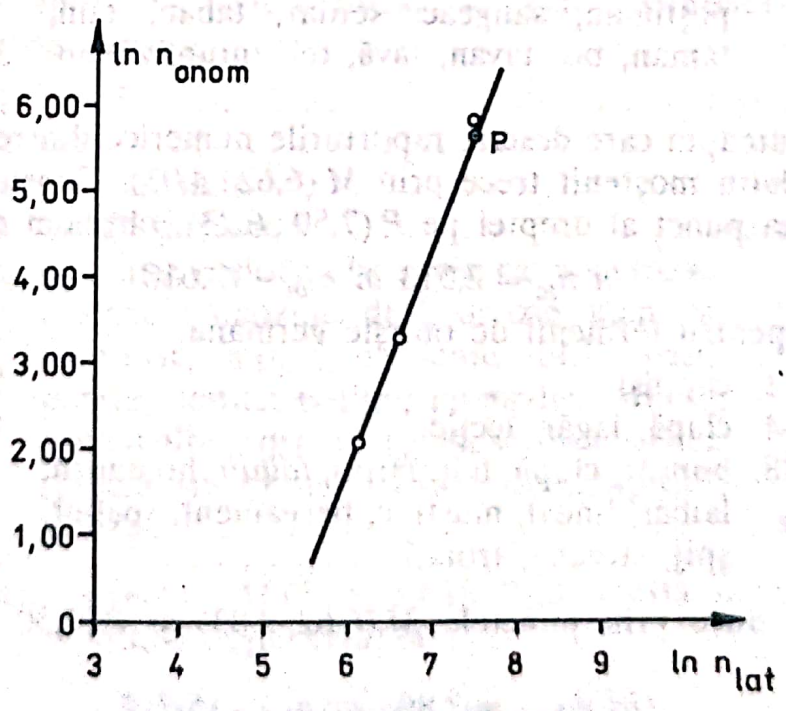
| $s \geq$ | | 1 | 3 | 4 |
|---|---|---|---|---|
| CUVINTE-BAZĂ | $ln\ n_{lat}$ | 7,50 | 6,62 | 6,14 |
| | $ln\ n_{onom}$ | 5,84 | 3,33 | 2,08 |

Faţă de dreapta determinată de punctele 3 şi 4 (fig. 16.11), a cărei ecuaţie este:

$$ln\ n_{onom} = 2{,}586\ ln\ n_{lat} - 13{,}794 \qquad (16.15)$$

punctul $P$ (7,475 ; 5,673) se situează din nou foarte aproape (ecartul este de 2,4%).

FIG. 16.11






Scanned with OKEN Scanner

În lumina celor de mai sus, putem afirma că ipoteza potrivit căreia raportul cantitativ global dintre diferitele componente etimologice ale vocabularului determinat cu ajutorul ecuaţiilor productivităţii lexicale (raport „vizualizat" de punctul *P*) este valabil şi pentru ecuaţiile productivităţii semantice se confirmă pe deplin. Între altele, acest rezultat ne permite şi să depăşim obstacolul lipsei de date de bază în situaţiile, destul de numeroase, cînd informaţiile certe de care se dispune nu ne furnizează decît poziţia unui singur punct al dreptei ce urmează a fi construită. Folosindu-l pe *P* drept un al doilea punct al dreptei, ştim acum că nu vom greşi prea mult (cu maximum 3—4%).

Astfel, din clasamentul cuvintelor provenite din turceşte:

| | | |
|---|---|---|
| 1. | tacîm . . . . . | 6 sensuri |
| 2—4. | chef, chior, perdea . . . . . | 5 sensuri |
| 5—14. | berechet, *cătrănit, chiorî,* culă, divan, raia, *tătarcă,* tinichea, toi, uluc . . . . . | 4 sensuri |
| 15—66. | *băcănie,* beilic, cafegiu, caimac, calup, catran, cazan, cazma, *cătrăni, ceaprăzărie,* ceauş, cerdac, cheag, chehaia, chenar, cirac, ciubuc, *cîntări,* cîrmîz, conac, cusur, cutie, deliu, *dulgherie,* haide, hambar, haram, haz, lambă, leafă, marafet, mazil, meterez, mofluz, papuc, *părăluţă, pehlivan, pehlivănie,* perghel, peşchir, peştiman, sangeac, serdar, taban, tain, taman, tas, tavan, tavă, tel, tutun, zăbun | 3 sensuri |

rezultă că dreapta care descrie raporturile numerice dintre elementele turcesc şi latin moştenit trece prin *M* (6,62 ; 4,03). Considerînd drept un al doilea punct al dreptei pe *P* (7,50 ; 6,23), obţinem ecuaţia:

$$ln\ n_{tc} = 2{,}513\ ln\ n_{lat} - 12{,}610 \qquad (16.16)$$

La fel, pentru termenii de obîrşie germană:

| | | |
|---|---|---|
| 1. | căptuşi . . . . . | 5 sensuri |
| 2—4. | clapă, lagăr, lecţie. . . . . | 4 sensuri |
| 5—18. | buncăr, clupă, falţ, fălţui, *fălţuit,* holendru, laibăr, lineal, muştiuc, pergament, spalier, şpiţ, streang, troacă . . . . . | 3 sensuri |

dreapta se duce prin punctele *M* (6,62; 2,83) şi *P* (7,50; 5,30) şi are ecuaţia:

$$ln\ n_{germ} = 2{,}803\ ln\ n_{lat} - 15{,}722 \qquad (16.17)$$



Scanned with OKEN Scanner

Pentru împrumuturile sîrbocroate:

1—2. posadă, priboi . . . . . 4 sensuri
3—9. cîrcel, cîrd, drug, pleoşti, *răzuş,* scoc, scai. 3 sensuri

dacă se ţine seama şi de cele 31 de cuvinte-bază care posedă cîte două sensuri şi anume: bălăcări, chiti 1, chiti 2, cujbă, dăinui, doniţă, dospi,drăgănele, gîrlici, izină, oraniţă, pernă, pîrpără, plută, postavă, priod, pripon, puhav, pup, raz, seci, sfor, sîrb, sloven, snagă,stative, strujniţă, şiţi, şoacăţ, vălug, vîrstă (vrîstă), dreapta poate fi trasată prin punctele *M* (7,05 ; 3,66) şi *P* (7,50 ; 4,69) şi ecuaţia ei are forma:

$$ln\ n_{scr} = 2{,}260\ ln\ n_{lat} - 12{,}266 \qquad (16.18)$$

Similar, pentru elementul ucrainian:

1. domol . . . . . 6 sensuri
2—8. calic, *calici,* crainic, oboroc, tarniţă, temnic, *toropeală* . . . . . 3 sensuri

adăugînd cele 25 de cuvinte-bază cu cîte două înţelesuri (buhai, cahlă clăpăug, cotigă, cuhnie, cuşcă, cuşmă, dubi, gotcă, hîrcă, hîtru, holercă, horn, nagodă, năvoloacă, năzări, ometiţă, pajură, postoroncă, prelucă, toloacă, tologi, toval, vidmă, zăhăi) se obţine punctul *M* (7,05 ; 3,43). Ecuaţia dreptei duse prin acesta şi prin punctul *P* (7,50 ; 5,05) este:

$$ln\ n_{ucr} = 3{,}580\ ln\ n_{lat} - 21{,}807 \qquad (16.19)$$

În cazul aportului lexical rusesc, ierarhia cuvintelor-bază după numărul lor de sensuri:

1. poştă . . . . . 6 sensuri
2. ceai . . . . . 5 sensuri
3—4. cavaler, talie . . . . . 4 sensuri
5—6. bantă, poştalion . . . . . 3 sensuri
7—38. balamuc, bielorus, bolşevic, calmuc, cazac, ciorpac, colhoznic, delă, divizie, manişcă, menşevic, mojic, narodnic, ofili, paşli, pensie, petliţă, poliţie, pristavlisi, prociti, proletcultist, prusac, rublă, scripcă, soviet, sprafcă, stroi, tract, uzbec, veleat, vipuşcă, zăloagă . . . . . 2 sensuri

ne furnizează punctul *M* (7,05 ; 3,64). Prin acesta şi prin punctul *P* (7,50 ; 4,95) se construieşte dreapta:

$$ln\ n_{rus} = 2{,}820\ ln\ n_{lat} - 16{,}233 \qquad (16.20)$$



Scanned with OKEN Scanner

Se observă că, la fel cum am procedat şi în studiul productivităţii lexicale, am lăsat deoparte la inventarierea termenilor de origine ucraineană şi rusească acele cuvinte cărora dicţionarele (DLRM coroborat cu DEX) le atribuie o dublă etimologie, fie aceasta ruso-ucraineană, ruso-poloneză, sau polono-ucraineană. Împreună cu nu prea numeroasele unităţi provenite numai din limba polonă, aceste cuvinte formează colecţia eteroclită grupată anterior sub genericul „nord-slav-medieval". Dintre ele, următoarele cuvinte-bază posedă minimum 2 sensuri distincte:

| | | |
|---|---|---|
| 1—2. | ba, podină . . . . . | 4 sensuri |
| 3—7. | bună, cneaz, ofiţer, şanţ, şcoală . . . . . | 3 sensuri |
| 8—34. | arendă, ataman, chilnă, darabană, drob, ghint, hatman, hoardă, jaf, muscal, palancă, pavăză, pisar, polcovnic, puşcă, schijă, scumpie, sponcă, stănog, şină, şirincă, şleahtă, şofran, toană, tovarăş, vajnic, zăton . . . . . | 2 sensuri. |

Lista de mai sus defineşte punctul $M$ (7,05 ; 3,53). Dreapta dusă prin el şi prin punctul $P$ (7,50 ; 4,02) are ecuaţia:

$$ln\, n_{n.s.m.} = 1{,}100\ ln\, n_{lat} - 4{,}228 \qquad (16.21)$$

care descrie raporturile materialului lexical zis „nord-slav medieval" cu fondul latin, din punctul de vedere al bogăţiei de sensuri.

După cum s-a putut vedea, mulţimile de cuvinte pe care ne-am bazat în elaborarea ecuaţiilor precedente au fost progresiv tot mai restrînse. Un „record" negativ în această privinţă îl reprezintă situaţia unităţilor lexicale formate în româneşte din elemente de origine latină. Pînă la nivelul de polisemie 3, nu întîlnim în DLRM decît următoarele cuvinte-bază din această categorie:

| | | |
|---|---|---|
| 1—2. | dacă, dintru . . . . . | 9 sensuri |
| 3. | ceva . . . . . | 6 sensuri |
| 4—7. | amnar, binecuvîntat, înapoi, mulţumi . . | 4 sensuri |
| 8—19. | aminti, alde, binecuvînta, ci, cică, dalb, deodată, depărta, încolo, înspre, odată, unealtă . . . . . | 3 sensuri |

Dreapta trasată prin punctul $M$ (6,62 ; 2,94) furnizat de această listă şi prin punctul $P$ (7,50 ; 4,75) are ecuaţia:

$$ln\, n_{rom(lat)} = 2{,}062\ ln_{lat} - 10{,}713 \qquad (16.22)$$



Scanned with OKEN Scanner

Gradul de precizie a rezultatelor ce se vor obţine cu ajutorul acestei ecuaţii va avea, fireşte, de suferit de pe urma insuficienţei datelor de bază, situaţie valabilă, mutatis mutandis, şi pentru unele dintre relaţiile anterioare.

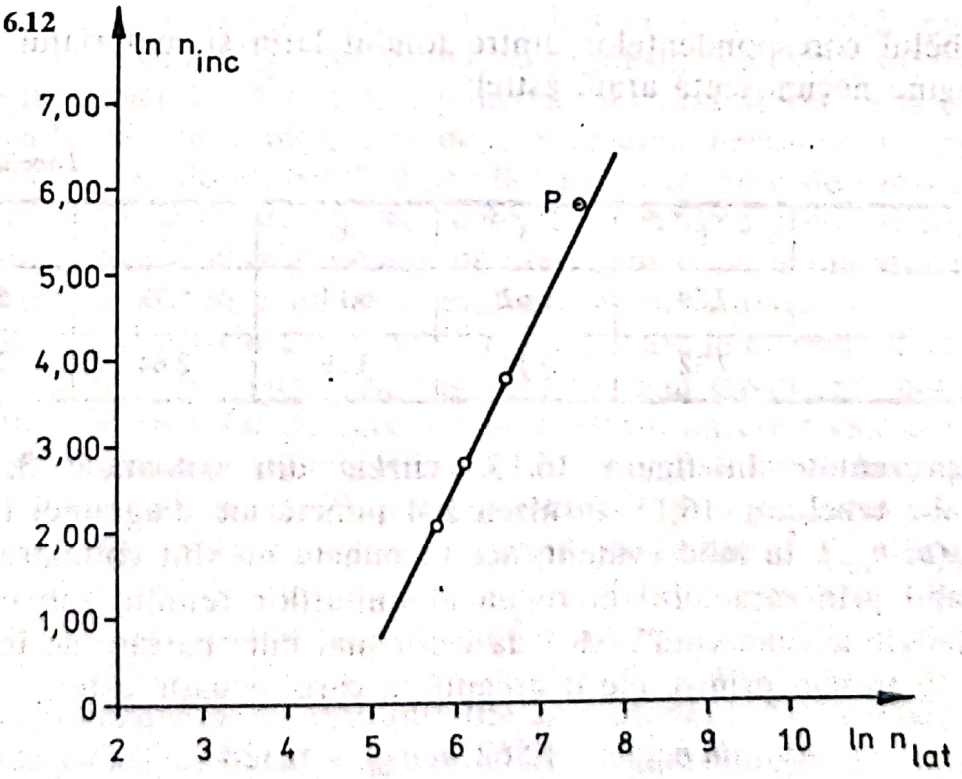
Ne-a mai rămas de examinat materialul eteroclit de origine incertă şi necunoscută. Cuvintele din prima categorie permit o trasare comodă şi destul de sigură a dreptei caracteristice, dat fiind că cele 3 puncte de coordonate (6,62; 3,78), (6,14; 2,83) şi (5,78; 2,08) rezultate din clasificarea termenilor-matcă:

| | | |
|---|---|---|
| 1—2. | cîrlig, talpă . . . . . | 7 sensuri |
| 3. | pop . . . . . | 6 sensuri |
| 4—7. | ciupi, stropşi, taler, trînti . . . . . | 5 sensuri |
| 8—16. | ciocni, cucuruz, găsi, gurgui, ia, pitic, tăvăli, tîrnă, ţeapă . . . . . | 4 sensuri |
| 17—43. | bui, bumb, burlan, ciot, comarnic, hîrb, înţepa, jumuli, mirodenie, murui, năbuşi, obîrşie, opinti, pestriţ, posomorî, pripi, sovon, ştiubei, şuviţă, teică, treaglă, trînd, ţarină, zăgaz, zănoagă, zgripţor, zimţ . . | 3 sensuri |

sînt aproape perfect coliniare (fig.16.12). Dreapta determinată de ele are ecuaţia:

$$\ln n_{inc} = 2.000 \ln n_{lat} - 9{,}463 \qquad (16.23)$$

FIG. 16.12






Scanned with OKEN Scanner

şi trece destul de aproape de punctul $P(7{,}50\,;\,5{,}79)$ rezultat din modelul productivităţii (abaterea ordonatei faţă de valoarea calculată este de cca. 4,5%, în fond remarcabil de mică pentru o colecţie de cuvinte atît de pestriţă).

Potrivit criteriului numărului de înţelesuri, cuvintele nederivate cu etimologie necunoscută prezintă următoarea ierarhie:

| | | |
|---|---|---|
| 1. | vorbă . . . . . | 11 sensuri |
| 2—3. | băga, străin . . . . . | 9 sensuri |
| 4—5. | butuc, moţ . . . . . | 7 sensuri |
| 6—9. | burduf, dar, pînză, stîrni . . . . . | 6 sensuri |
| 10—14. | baie, mălai, melc, pîlnie, scormoni . . . . | 5 sensuri |
| 15—22. | aprig, argea, băiat, cruţa, gogoaşă, pociumb, solz, zbîrli . . . . . | 4 sensuri |
| 23—63. | aromi, bănuţ, băţ, breabăn, brînă, clănţăni, cotărit, cotor, dibui, fofelniţă, fulg, ghin, janţ, momîie, noian, oblon, ortoman, pisc, plisc, pocie, pojghiţă, piţărău, prepeleac, pungă, răbda, răsfug, răşlui, răvar, reteza, scociorî, sfîrlează, stăpîn, titirez, ţipa 1, ţipa 2, ţîţînă, ţurţur, văgăună, zborşi, zgău, zgîndări . . . . . | 3 sensuri |

Tabelul corespondenţelor dintre fondul latin şi materialul lexical de origine necunoscută arată astfel:

*Tabelul 16.15.*

| $s \geq$ | 1 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $ln\ n_{lat}$ | 7,50 | 6,62 | 6,14 | 5,78 | 5,40 |
| $ln\ n_{nec}$ | 7,32 | 4,14 | 3,09 | 2.64 | 2,20 |

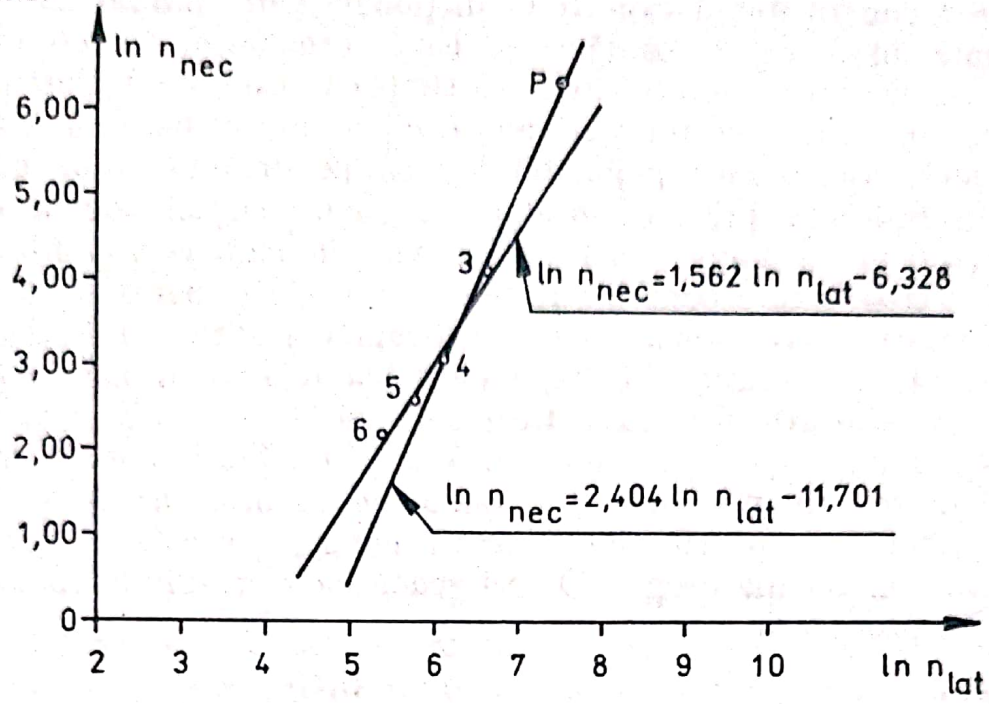
Reprezentate în figura 16.13., cifrele din coloanele 3, 4, 5, şi 6 ale tabelului 16.15. furnizează 4 puncte ale diagramei funcţiei $ln\ n_{nec}(ln\ n_{lat})$. În mod evident, aceste puncte nu sînt coliniare, lucru explicabil prin caracterul eterogen al unităţilor reunite sub eticheta „etimologie necunoscută". Metoda celor mai mici pătrate ne îngăduie totuşi să trasăm printre ele o dreaptă, a cărei ecuaţie este:

$$ln\ n_{nec} = 1{,}562\ ln\ n_{lat} - 6{,}328 \qquad (16.24)$$



Scanned with OKEN Scanner

FIG. 16.13

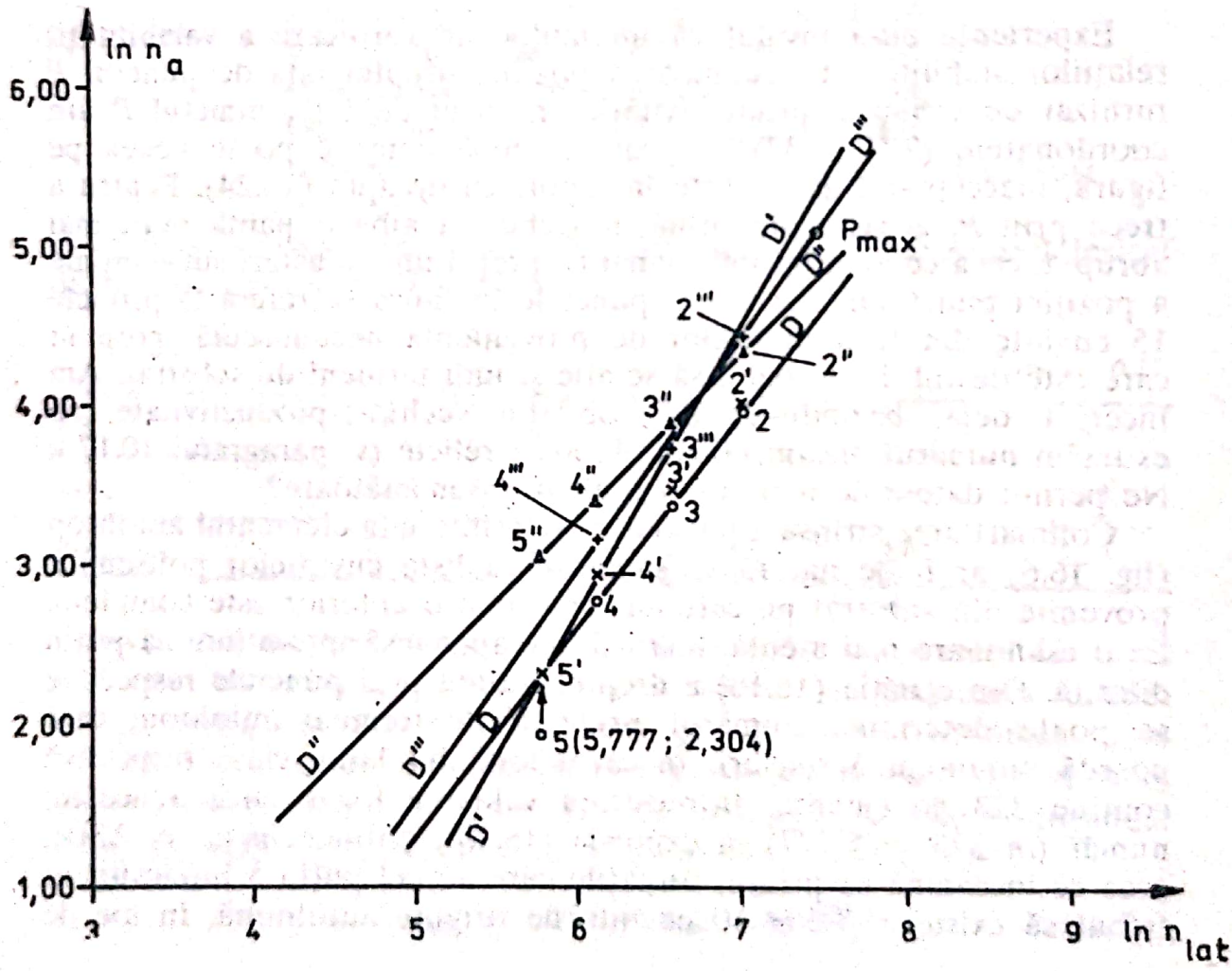



Experienţa ne-a învăţat că un mijloc de verificare a valabilităţii relaţiilor stabilite este examinarea poziţiei dreptei faţă de punctul *P* furnizat de ecuaţiile productivităţii. În cazul de faţă, punctul *P* are coordonatele (7,50 ; 6,37) şi e situat, după cum se poate vedea pe figură, inacceptabil de departe în raport cu dreapta (16.24). Pentru a trece prin *P*, aceasta din urmă ar trebui să aibă o pantă mult mai abruptă, ceea ce ar fi posibil numai cu preţul unei ajustări substanţiale a poziţiei punctelor 5 şi 6. Or, punctele în cauză se referă la primele 15 cuvinte din lista unităţilor de provenienţă necunoscută, grup în care este destul de probabil să se afle şi unii termeni de substrat. Am încercat deja, bazîndu-ne pe corelaţia vechime-productivitate, să estimăm numărul maxim posibil de atari relicte (v. paragraful 10.17.). Ne permit datele de mai sus o evaluare asemănătoare?

Coliniaritatea strînsă a punctelor referitoare la elementul autohton (fig. 16.6) ar fi de natură să sugereze că lista cuvintelor polisemice provenite din substrat pe care am prezentat-o anterior este completă. La o examinare mai atentă, lucrurile ne apar însă într-o lumină puţin diferită. Din ecuaţia (16.10) a dreptei trasate prin punctele respective se poate determina numărul probabil de termeni autohtoni care posedă minimum 5 sensuri. În cazul fondului latin, clasa respectivă conţine 323 de cuvinte. Introducînd valoarea logaritmică a acestui număr (*ln* 323 = 5,777) în expresia (16.10), obţinem $ln\, n_a = 2{,}304$, ceea ce înseamnă că printre unităţile care au cel puţin 5 înţelesuri ar trebui să existe $e^{2{,}304} = 10$ cuvinte de origine autohtonă, în loc de

numai 7 cîte figurează explicit în dicţionar. Pare natural ca cele 3 elemente lipsă să le căutăm în lista termenilor cu etimologie necunoscută, dar se observă lesne că simpla transferare a 3 asemenea cuvinte în lista substratului nu rezolvă problema. Ea asigură într-adevăr ascensiunea punctului 5 pînă pe dreapta D de ecuaţie (16.10), însă din pricina faptului că, odată, cu extinderea clasei cuvintelor cu cel puţin 5 sensuri, sporesc în mod corespunzător şi celelalte clase care o conţin, punctele 4, 3 şi 2 părăsesc dreapta într-o mişcare ascensională generală, ocupînd poziţii noi 4′, 3′ şi 2′ (fig. 16.14). În această situaţie, liniaritatea relaţiei ar mai putea fi salvată în felul următor: noua dreaptă D′ să fie trasată prin punctele 5′ şi 4′, aducerea pe D′ a punctelor 3′ şi 2′ rămase la cote inferioare realizîndu-se prin preluarea din fondul de cuvinte „necunoscute" a exact atîţia termeni cît sînt necesari pentru respectarea condiţiei rezultate din ecuaţia dreptei D′. Să vedem la ce rezultate conduce o astfel de procedură.

FIG. 16.14



Coordonatele punctelor 5′ şi 4′ sînt, respectiv, (5,78 ; 2,30) şi (6,14 ; 2,94). Dreapta D′ pe care ele o determină are ecuaţia:

$$ln\, n_a = 1{,}777\ ln\, n_{lat} - 7{,}926 \qquad (16.25)$$

Punctele dreptei de abscise 6,62 şi 7.05, corespunzătoare mulţimilor de cuvinte care posedă minimum 3 şi, respectiv, 2 sensuri, vor avea ordonatele 3,84 şi 4,60, ceea ce revine la a spune că pentru respectarea liniarităţii ar trebui să existe $e^{3,84} = 47$ de cuvinte autohtone în prima categorie şi $e^{4,60} = 100$ în cea de a doua. Avînd în vedere că în lista substratului chiar după transferul celor 3 cuvinte fără etimon cunoscut, aceste două clase nu conţin decît 32 şi, respectiv, 55 de membri, devine necesară preluarea de la „necunoscute" a încă 15 cuvinte cu cîte 3 şi 30 de cuvinte cu cîte 2 înţelesuri. Lucrul n-ar fi în principiu imposibil, căci rezerva de termeni de origine încă neexplicată este suficient de bogată, dar el vine în contradicţie cu rezultatul furnizat de studiul productivităţii lexicale, potrivit căruia numărul total de cuvinte provenite din substrat nu poate depăşi limita de cca. 180. Or, din ecuaţia (16.25) deducem că punctul de abcisă 7,50 corespunzător totalităţii cuvintelor de origine latină din DLRM are ordonata 5,40 sau, cu alte cuvinte, că, acceptînd drept valabilă relaţia exprimată de dreapta D′, termenii autohtoni din dicţionar ar trebui să fie în număr de $e^{5,40} = 222$. Este evident că dreapta D′ e prea abruptă.

Fireşte, panta ei ar putea fi micşorată prin extinderea achiziţiilor din domeniul cuvintelor cu mai multe sensuri în detrimentul celor cu sensuri mai puţine. În această privinţă, un caz limită l-ar reprezenta transferarea în bloc la „autohtone" a tuturor celor 14 cuvinte de origine necunoscută care au mai mult de 4 sensuri. Noua dreaptă D" dusă prin punctele 5′′ (5,78 ; 3,04) şi 4" (6,14 ; 3,40) astfel obţinute, are ecuaţia:

$$ln\, n_a = ln\, n_{lat} - 2{,}740 \qquad (16.26)$$

În conformitate cu ea, numărul total de cuvinte de substrat ar trebui să fie de 117, ceea ce pare acceptabil, la prima vedere. Rămîne să mai calculăm însă ce achiziţii sînt necesare în grupele unităţilor cu 2 şi, respectiv, 3 sensuri ca nu cumva vreunul dintre aceste rezultate să se dovedească absurd ( o pantă prea mică a dreptei poate conduce la valori negative ale numărului de cuvinte ce trebuie transferate).

Introducînd în ecuaţia (16.26) valorile 6,62 şi 7,05 ale abscisei punctelor 3" şi 2", rezultă ordonatele 3,88 şi 4,31, adică un total de



Scanned with OKEN Scanner

$e^{3,88} = 48$ de cuvinte pentru clasa termenilor cu cel puţin 3 sensuri şi $e^{4,31} = 74$ pentru clasa celor cu minimum 2 sensuri.

Tabloul complet al preluărilor de elemente „necunoscute" în fondul autohton este atunci următorul:

— totalitatea celor 14 cuvinte cu cîte 5 şi mai multe înţelesuri;
— nici un cuvînt cu 4 sensuri;
— 5 cuvinte cu cîte 3 sensuri;
— 3 cuvinte cu cîte 2 sensuri;
— 3 cuvinte monosemantice.

După cum se observă, rezultatul nu este absurd, dar aceasta nu îl împiedică să fie suspect. Conform lui, cele mai bogate în sensuri cuvinte cu etimologie obscură ar trebui să provină toate din substrat, în timp ce din restul de nu mai puţin de 1493 de termeni numai 11(!) ar fi autohtoni. Chiar considerînd polisemia drept un indiciu de vechime, ceea ce e departe de a fi demonstrat, disproporţia este prea mare pentru a face plauzibil un asemenea scenariu. Ce-i de făcut?

Analizele anterioare ne-au arătat că un mod de a reduce riscul de eroare este trasarea dreptelor caracteristice în aşa fel încît să treacă prin punctul $P$ rezultat din ecuaţiile productivităţii lexicale ca raport numeric ideal dintre componentele etimologice considerate. În cazul substratului, am determinat (v. paragraful 10.17.), pe baza corelaţiei vechime-productivitate, ecuaţia variantei maximale a fondului autohton sub forma:

$$\ln n_{a.max.} = 5,07 - 0,23\ d$$

unde $d$ era numărul de derivate, pe care egalîndu-l cu zero găsim ordonata punctului $P$ în poziţia sa cea mai înaltă. $P$ are aşadar coordonatele (7,50 ; 5,07). Din condiţia banală ca numărul de cuvinte cu cîte minimum 4 sensuri să fie cel puţin egal cu numărul de cuvinte cu cîte minimum 5 sensuri, cumulată cu condiţia situării pe dreaptă a punctului $P$, se obţine ecuaţia:

$$\ln n_a = 1,413\ \ln n_{lat} - 5,526 \qquad (16.27)$$

reprezentată în figura 16.14. sub forma dreptei D''. Calculul ordonatelor punctelor 2'', 3'' şi 4'' ne conduce în cele din urmă la următorul tablou al transferurilor ce ar trebui operate între clasa cuvintelor cu etimologie necunocută şi substrat:

— 7 cuvinte cu 5 sau mai multe sensuri;
— nici un cuvînt cu 4 sensuri;



Scanned with OKEN Scanner

— 10 cuvinte cu cîte 3 sensuri;
— 15 cuvinte cu cîte 2 sensuri;
— 35 de cuvinte monosemantice.

Modificările propuse acum au un aspect mult mai „normal", mai echilibrat, decât precedentele. Singurul element de natură să ne pună pe gînduri este faptul că din grupa cuvintelor care posedă cîte 4 sensuri nici unul nu a fost reţinut ca posibil autohton, deşi cel puţin pentru substantivul *argea* (cf.v.maced., αργελλα, cimerian αργιλλα) explicaţia prin substrat, care a ispitit pe mulţi lingvişti (Haşdeu, Puşcariu, Densuşianu)[1], pare greu de combătut. Ecuaţia ar trebui deci corectată astfel încît să ţină seama de această precizare a provenienţei unui termen „tetravalent", dreapta continuînd să treacă prin punctul obligat $P(7{,}50\,;\,5{,}07)$. Noua formă a ecuaţiei este în acest caz, următoarea:

$$ln\, n_a = 1{,}337\ ln\, n_{lat} - 4{,}959 \qquad (16.27)$$

iar transferurile de cuvinte cu etimologie necunoscută către clasa „autohtonelor" sînt acum:

— 9 cuvinte cu cîte 5 şi mai multe sensuri;
— un cuvînt cu 4 sensuri;
— 10 cuvinte cu cîte 3 sensuri;
— 15 cuvinte cu cîte 2 sensuri;
— 32 de cuvinte monosemantice.

Se cuvine să rememorăm faptul că schimbarea intervenită nu a afectat clasele cuvintelor cu 2 şi 3 sensuri şi prea puţin pe cea a cuvintelor care nu au decît un singur înţeles. Considerînd lista transferurilor de mai sus drept cea mai verosimilă, să privim mai îndeaproape rezultatul referitor la unităţile lexicale ce cunosc minimum 5 accepţiuni diferite.

Din modul cum a fost determinat numărul lor rezultă că valoarea obţinută trebuie socotită drept minimă[2] Aşadar, dintre primele 15 elemente din lista cuvintelor de origine necunoscută, cel puţin 9 provin din substrat. Probabilitatea acestei apartenenţe trebuie însă diferenţiată şi ea pe clase de polisemie.

---

[1] Pentru bibliografia cazului, v. ILR II, 1969, p. 328.

[2] Încercarea de a o mai diminua încă ne aduce, cum s-a văzut, în situaţia de a nu mai putea accepta prezenţa nici unui termen de substrat printre cuvintele de origine obscură care posedă cîte 4 înţelesuri distincte.



Scanned with OKEN Scanner

Se calculează uşor că, în ipoteza respectării ecuaţiei (16.28) numărul unităţilor de substrat creditate cu cel puţin 6 sensuri trebuie să fie de cca. 9—10. Cum lista cuvintelor autohtone nu conţine decît 3 termeni de acest fel *(de, moş, ţarc)*, conchidem că restul de 6—7 se ascund printre următoarele 9 elemente de obîrşie neidentificată: *băga, burduf, butuc, dar, moţ, pînză, stîrni, străin, vorbă.* Pentru 3 dintre ele *(băga, butuc, pînză)* s-a propus deja explicaţia prin substrat[1] , iar *burduf* are, deşi DLRM şi DEX nu o spun, un echivalent[2] în alb. *burdhë* „sac". De subliniat că *butuc* şi *pînză* se numărau şi printre termenii recomandaţi drept autohtoni de modelul productivităţii lexicale (v. paragraful 10.17.). Care sînt celelalte 2—3 cuvinte de substrat din grupul rămas?

Nu riscăm un răspuns, dar nu ne putem reţine să observăm că un cuvînt socotit îndeobşte drept slav[3] ca *vorbă* (cf. v. sl. *dvorĭba*) figurează printre unităţile ce ar putea proveni din fondul autohton atît în lista stabilită pe baza modelului $M_p$ cît şi în cea furnizată de studiul polisemiei. Un argument în plus pentru originea autohtonă a lui vorbă poate fi extras din examinarea membrului drept al relaţiei (16.28). Faptul că acesta devine egal cu zero pentru o valoare a argumentului $ln\ n_{lat} = 3{,}709$ înseamnă că în clasificarea cuvintelor româneşti după gradul lor de polisemie primul termen de substrat apare de-abia după ocurenţa a $e^{3{,}709} = 41$ de elemente de origine latină. Or, primele 40 de unităţi de acest fel au toate cîte cel puţin 13 înţelesuri diferite. Este deci de aşteptat ca cel mai bogat în sensuri cuvînt de origine autohtonă să posede 13 înţelesuri, condiţie îndeplinită întocmai de conjuncţia *de*. Pentru a afla şi numărul probabil de sensuri ale celui de al doilea termen de substrat, ajunge să egalăm acelaşi membru drept (16.28) cu $ln\ 2 = 0{,}693$. În final găsim că secundantul lui *de* ar trebui să posede 10—11 înţelesuri, ceea ce se întîmplă numai cu substantivul *vorbă,* următoarele clasate în lista cuvintelor de provenienţă necunoscută *(băga, străin*[4]) numărînd doar 9 sensuri. Lucrurile care păreau să se clarifice se complică însă din nou dacă supunem unui examen asemănător ecuaţia (16.8) ce descrie raporturile

---

[1] Tustrele figurează în lista cuvintelor autohtone care lipsesc din albaneză propusă de I.I. Rusu (*op. cit.*, p. 215—216).

[2] ILR II. p. 329.

[3] V. de exemplu A. Rosetti ILR III, ed. V, 1964, p. 105

[4] *Străin* are de altfel mari şanse de a fi moştenit din latineşte, cum a arătat Mioara Avram într-o comunicare citată de I. Coteanu şi M. Sala *(Etimologia si limba română.* Ed. Academiei RSR, 1987, p. 109)



Scanned with OKEN Scanner

dintre elementul vechi slav şi fondul primordial latin. Anularea membrului al doilea conduce şi în acest caz la acelaşi număr 13 pentru sensurile celui mai „polisemantic" cuvînt slav, cu toate că nu există termeni vechi slavi recunoscuţi ca atare[1] în DLRM şi DEX care să cunoască mai mult de 9 sensuri (ca şi în clasa unităţilor de origine necunoscută, numărul acestora este egal cu 2; e vorba de *sfînt* şi *val*). Elementul vechi slav este prin urmare şi el îndreptăţit să şi-l apropie pe *vorbă*. Ba chiar nevoia sa de cuvinte cu grad ridicat de polisemie e mai acută decît cea a substratului, care dispune oricum de un termen cu 13 înţelesuri. Pe de altă parte însă, cooptarea lui *vorbă* în grupul cuvintelor de obîrşie slavă veche riscă să dezechilibreze, din pricina numărului ridicat de derivate ale acestuia (17), balanţa productivităţii, periclitînd liniaritatea ecuaţiei caracteristice. Speculaţiile noastre nu oferă, după cum se vede, o soluţie, dar încearcă să evidenţieze posibilitatea de a gîndi problemele diagnozei etimologice într-un mod radical diferit de cel de pînă acum.

Referitor la al doilea cuvînt susceptibil de a fi atribuit substratului din grupul *dar, moţ, stîrni, străin,* preferinţele noastre ar merge către primul termen, în ciuda improductivităţii sale[2] (fenomen firesc, întrucît instrumentele gramaticale sînt, prin natura lor, predispuse la sterilitate). Ne bazăm opţiunea pe constatarea că nici una dintre conjuncţiile româneşti frecvent[3] folosite nu este împrumutată. Limba latină şi substratul (cazul; conjuncţiei *de*) constituie sursele exclusive ale acestui tip de conectori.

Alte vreo 2—3 cuvinte autohtone se ascund probabil în grupul *baie, mălai, melc, scormoni.* Criteriul productivităţii, care nu este niciodată decisiv, dar posedă o certă valoare orientativă, le-ar recomanda pe *mălai* şi *scormoni* (ambele au în DLRM cîte 5 derivate).

Să mai notăm, în fine, că nici ipoteza potrivit căreia elementele de substrat necomune cu albaneza s-ar afla toate printre unităţile zise „cu etimologie necunoscută" nu este inatacabilă. Adjectivele *mic* şi *mare,* socotite îndeobşte ca aparţinînd fondului latin, figurează şi ele printre candidaţii la indigenat[4], iar gradul lor remarcabil de polisemie (ambele sînt prezente în DLRM cu cîte 8 sensuri distincte) este de

---

[1] Dintre unităţile cu etimologie dublă bulgaro-sîrbocroată substantivul *colţ* figurează în DLRM cu 11 sensuri diferite.

[2] Nu absolute, odată ce *dar* a dat totuşi naştere formaţiei *dar(ă)mite.*

[3] Am adăugat complementul „frecvent" ca o justificare a „păcatului" de a fi utilizat noi înşine, cu numai cîteva rînduri mai sus, franţuzismul *or.*

[4] v. I.L.R. II, 1969, pp. 332 şi 344.



Scanned with OKEN Scanner

natură să contribuie la modificarea imaginii pe care ne-am format-o mai sus în legătură cu tabloul transferurilor către clasa cuvintelor autohtone.

Revenind la obiectivul nostru imediat, care era stabilirea unei ecuaţii vizînd raporturilor dintre materialul lexical de origine necunoscută şi fondul latin moştenit, ne vom declara mulţumiţi deocamdată cu poziţionarea dreptei caracteristice astfel încît să treacă prin punctele 4, 3 şi *P* (fig. 16.13.). Coliniaritatea acestor puncte constituie o garanţie că, procedînd astfel, greşim mai puţin decît în cazul oricărei alte opţiuni şi, oricum, nu atît de mult cît am făcut-o probabil la trasarea dreptelor corespunzătoare unor aporturi mai sărace, cînd nu cunoşteam cu certitudine decît poziţia unui singur punct. Ecuaţia rezultantă are forma:

$$ln\ n_{nec} = 2{,}404\ ln\ n_{lat} - 11{,}701 \qquad (16.29)$$

Cu aceasta, modelul $M_s$ al organizării polisemice a vocabularului a ajuns la fel de complet, dacă nu şi la fel de precis, ca modelul $M_p$ al productivităţii lexicale. Rămîne să îl traducem în valori numerice în felul în care am făcut-o şi pentru $M_p$ (v. capitolul 15) şi să tragem concluziile ce se impun din confruntarea rezultatelor.

## 17. Verificări şi corecţii

Avînd în vedere aproximaţiile la care ne-am văzut obligaţi să recurgem, din lipsă de date, apreciem că o verificare suplimentară a concordanţei ecuaţiilor cu realitatea lingvistică modelată nu poate fi decît binevenită. Toate relaţiile stabilite erau de forma:

$$ln\ n_i = a_i\ ln\ n_{lat} - b_i$$

Or, ştim deja, din analiza ecuaţiilor omomorfe ale productivităţii lexicale, că, prin anularea membrului drept, se obţine o valoare particulară a lui $n_{lat}$ avînd semnificaţia următoare: numărul respectiv, $n_{lat} = exp\ (b_i / a_i)$ reprezintă numărul de cuvinte de origine latină care posedă cel puţin tot atîtea înţelesuri ca şi cel mai bogat în sensuri cuvînt de originea *i*. Cunoscînd valoarea acestui parametru, putem



Scanned with OKEN Scanner

determina, folosindu-ne fie de tabelul 16.3., fie de graficul din fig. 16.1 (curba B), cîte sensuri ar trebuie să aibă, conform ecuaţiilor modelului $M_s$, cel mai „polisemic" termen provenit din limba *i*. Nu are rost să încărcăm lucrarea cu toate calculele reclamate de un atare demers, dar prezentăm totuşi un exemplu spre edificarea cititorului.

Ce nivel maxim de polisemie este probabil să fi atins împrumuturile din limba germană? Corelaţia *german-latin moştenit* e dată de expresia (16.16), stabilită cu ajutorul unui singur punct şi a ipotezei, postulate prin analogie cu situaţia altor aporturi, că dreapta caracteristică trece prin punctul $P\,(7{,}50\,;\,5{,}30)$, ce ne este furnizat de modelul productivităţii lexicale. Egalînd cu zero membrul drept al ecuaţiei:

$$ln\ n_{germ} = 2{,}803\ ln\ n_{lat} - 15{,}722$$

obţinem imediat $ln\ n_{lat} = 5{,}609$. Această valoare corespunde în tabelul 16.3. intervalului $s = 5 \div 6$, ceea ce înseamnă că cel mai polisemic cuvînt de sorginte germană ar trebui să posede 5 sau 6 înţelesuri. Presupunerea este confirmată de prezenţa în fruntea listei împrumuturilor germane a verbului *căptuşit,* căruia DLRM îi menţionează 5 sensuri distincte. Este şi aceasta o dovadă că ecuaţia (16.16) modelează destul de bine fenomenul, în ciuda artificiului de care am făcut uz la elaborarea ei.

Tabelul 17.1. prezintă comparativ numărul maxim de sensuri la care ar fi trebuit să acceadă, potrivit modelului matematic, cuvintele de diferite orgini faţă cu numărul maxim efectiv de înţelesuri consemnate de dicţionar.

*Tabelul 17.1*

| | SLAV | SUBSTR. | „FR." | LAT.SAV. | BG. & SCR. | BG. | MAGH. | INC. | NEC. | TC. |
|---|---|---|---|---|---|---|---|---|---|---|
| $s_{max.calc.}$ | 13 | 11—12 | 12 | 11 | 10 | 9 | 8—9 | 8 | 8 | 7 |
| $s_{max.real}$ | 9 | 13 | 9 | 13 | 11 | 7 | 9 | 7 | 11 | 6 |

| | R(LAT) | GREC | ONOM. | SCR. | GERM. | ITAL. | N.S.M. | RUS | UCR |
|---|---|---|---|---|---|---|---|---|---|
| $s_{max.calc.}$ | 6—7 | 6—7 | 6—7 | 6 | 5—6 | 5—6 | 12 | 5 | 4 |
| $s_{max.real}$ | 9 | 7 | 8 | 4 | 5 | 6 | 4 | 6 | 6 |

După cum se vede, în general, se înregistrează o bună concordanţă între valorile calculate şi cele reale. Considerînd că ecarturi de numai o unitate între cele două şiruri de cifre din tabelul 17.1, situaţie ce



Scanned with OKEN Scanner

caracterizează componentele bulgaro-sîrbocroată, maghiară, incertă, turcică, grecească, germană, italiană, rusă, constituie abateri absolut acceptabile şi că deci aporturile respective pot fi socotite drept corect modelate de ecuaţiile noastre, să examinăm pe rînd cazurile în care diferenţele ating valori mai mari.

Un ecart considerabil, ce se cere explicat, apare în cazul componentei slave vechi. Trasată mai exact decât toate celelalte corelaţii grafice (cu sigura excepţie a legăturii „francez" — latin, reprezentată tot cu ajutorul a 5 puncte), dreapta de ecuaţie (16.8) prezintă, graţie bogăţiei relative a datelor de bază, o credibilitate suficientă pentru ca explicaţia nepotrivirii dintre rezultatul calculului şi situaţia reală (13 sensuri expectate, faţă de numai 9 înregistrate de cele mai „polisemice" cuvinte slave: *sfânt* şi *val*) să nu o căutăm în aspectul ecuaţiei, ci în componenţa mulţimii de cuvinte ce a servit la construirea acesteia. Cu alte cuvinte, mai degrabă decît să punem la îndoială corectitudinea relaţiei matematice utilizate, este cazul să cercetăm dacă nu cumva în zona termenilor cu multe sensuri s-au produs omisiuni semnificative. La întrebarea ce alte cuvinte cu număr mare de înţelesuri ar mai putea fi vechi slave, răspund „prezent" două unităţi lexicale, una căreia DLRM şi DEX îi atribuie o etimologie dublă bulgaro-sîrbocroată şi anume cuvîntul *colţ* şi o alta, de origine obscură, substantivul *vorbă*. Ambele figurează în DLRM cu un evantai de cîte 11 înţelesuri diferite. Or, ecuaţia (16.8) pe care o transcriem din nou aici:

$$ln\ n_{sl} = 1{,}860\ ln\ n_{lat} - 6{,}832$$

ne arată că prezenţa a două cuvinte cu 11 sensuri fiecare este perfect echivalentă cu prezenţa unui singur cuvînt cu 13 sensuri.

Într-adevăr, fondul latin moştenit conţine 60 de termeni cu cîte 11 înţelesuri. Introdusă în ecuaţia (16.8), valoarea $ln\ 60 = 4{,}094$ ne furnizează $n_{sl} \approx 2$, ceea ce arată că simpla transferare în clasa împrumuturilor vechi slave a celor două cuvinte menţionate e suficientă pentru rezolvarea aparentei contradicţii dintre situaţia reală şi cea calculată. Este de reţinut faptul că, după ce în paragraful precedent se conturaseră unele argumente în favoarea provenienţei din substrat a controversatului cuvînt *vorbă*, acum balanţa s-a aplecat din nou, şi destul de energic, către ipoteza originii vechi slave.

Alte nepotriviri evidente privesc aporturile „francez" şi latin savant. Le discutăm împreună, fiindcă, pe de o parte, se ştie că cele două categorii de cuvinte sînt greu discernabile, rocadele de termeni dintr-una într-alta fiind oricînd posibile, şi, pe de altă parte, că ele

prezintă situaţii complementare. Exact ceea ce lipseşte elementului „francez", şi anume două cuvinte cu cîte cel puţin 11 înţelesuri, posedă în exces elementul latin savant. Totul s-ar rezolva prin transferarea în clasa neologismelor „franceze" a cuvintelor *linie* (13 sensuri şi formă perfect compatibilă cu provenienţa din franceză) şi *punct* (11 sensuri, dar aspect latin). Forma acestui al doilea cuvînt se opune explicaţiei prin franceză, însă nu trebuie să uităm că majoritatea celor 11 sensuri ale sale sînt totuşi preluate din franţuzeşte. Din acest punct de vedere, el s-ar încadra perfect în categoria, bogat ilustrată în DLRM, a cuvintelor considerate drept provenind din franceză exclusiv pe considerente semantice, aspectul lor contrazicînd, în parte sau total, o atare filiaţie *(dispune, introduce, libertate, practică, rectifica* etc.).

Tot împreună putem trata cazurile componentelor bulgară şi sîrbocroată ale vocabularului românesc, care au drept trăsătură comună o diferenţă de două unităţi între polisemia maximă anticipată prin calcul şi cea reală. Rezerva de termeni constituită de mulţimea unităţilor cu dublă etimologie bulgaro-sîrbocroată oferă posibilităţi de diminuare a acestei, de altfel puţin însemnate, nepotriviri.

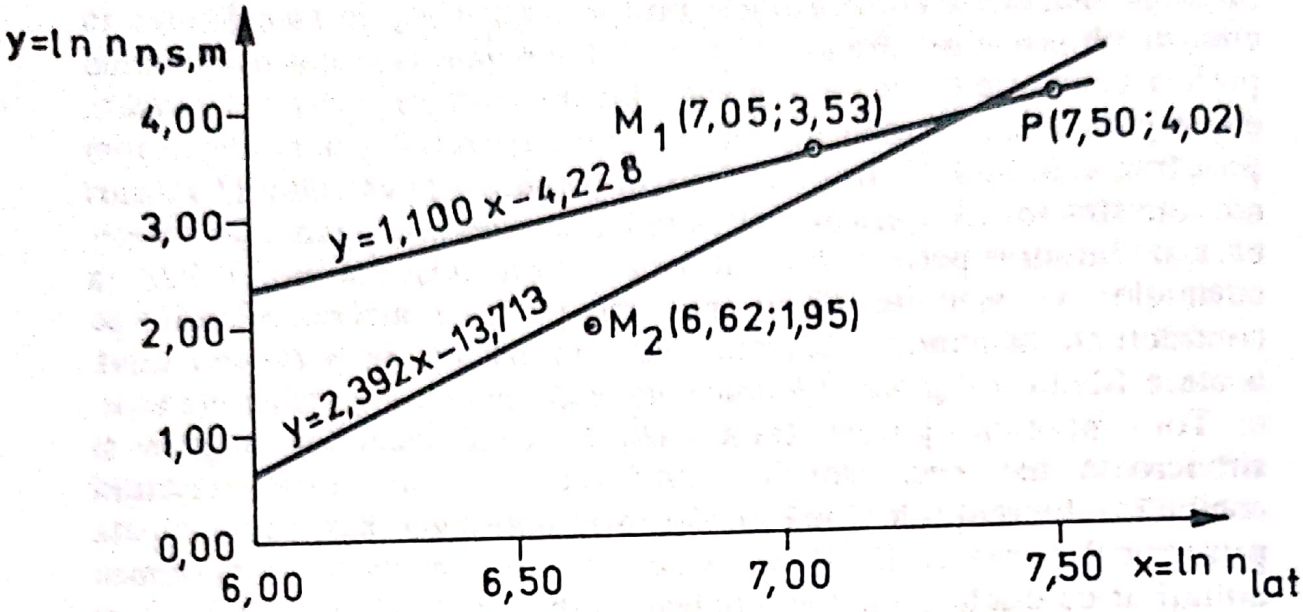
O diferenţă mai mare survine în cazul cuvintelor de origine necunoscută, dar şi aici lucrurile se rezolvă lesne prin îndepărtarea din listă a autohtonului *băga* şi a probabil slavului *vorbă,* care îşi au locul în altă parte. Dacă observăm că, în raport cu restul cuvintelor de obîrşie ucraineană, care nu depăşesc niciodată 3 sensuri, adjectivul *domol,* cu cele 6 înţelesuri ale sale, reprezintă o simplă excepţie, nu ne vom arăta îngrijoraţi nici de abaterea semnalată de ultima coloană a tabelului 17.1. Tot ca excepţii trebuie privite şi fruntaşele listei compuselor româneşti din elemente latine *dacă* şi *dintru,* care posedă cîte 9 sensuri distincte, cu 3 mai multe decît secondantele lor imediate. Ele ocupă o poziţie aparte şi din pricina faptului că, în cazul lor, nu este vorba de sensuri lexicale, ci gramaticale, adică de tipuri de raporturi sintactice pe care le pot exprima cei doi conectori. Atari situaţii atipice nu sînt oricum de natură să ne forţeze la o revizuire a ecuaţiilor stabilite.

În cele din urmă, singura contradicţie ireductibilă rămîne cea privitoare la aportul zis „nord-slav medieval". Ecuaţia (16.21) prognoza existenţa unui cuvînt cu 12 sensuri avînd etimologie multiplă nord-slavă, cînd în realitate cel mai „polisemic" termen de această provenienţă nu depăşeşte 4 înţelesuri. E limpede că relaţia matematică folosită nu este conformă cu realitatea. Se poate, de altfel, observa în figura 17.1 că dreapta de ecuaţie (16.21) trece foarte departe de punctul de coordonate (6,61; 1,95) care corespunde grupului de cuvinte



Scanned with OKEN Scanner

FIG. 17.1




cu cîte 3 şi 4 sensuri, punct pe care noi nu l-am luat în considerare, în baza convenţiei de a neglija mulţimile de cuvinte alcătuite din mai puţin de 10 elemente. Derogînd acum de la această constrîngere liber asumată, să trasăm dreapta corectată printre punctele $P$ (7,50; 4,02), $M_1$(7,05; 3,53) şi $M_2$(6,62; 1,95). Ecuaţia ce rezultă este:

$$ln\ n_{n.s.m.} = 2{,}392\ ln\ n_{lat} - 13{,}713 \qquad (17.1)$$

şi, în conformitate cu ea, cuvintele din clasa *n.s.m.* nu mai pot depăşi 5 sensuri, ceea ce concordă întrutotul cu realitatea.

În urma acestei rectificări, modelul $M_s$ capătă înfăţişarea de mai jos:

$n_{sl} = exp\ (1{,}860\ ln\ n_{lat} - 6{,}832)$
$n_a = exp\ (1{,}282\ ln\ n_{lat} - 5{,}102)$
$n_{fr} = exp\ (2{,}361\ ln\ n_{lat} - 9{,}081)$
$n_{lat.sav} = exp\ (1{,}791\ ln\ n_{lat} - 7{,}246)$
$n_{bg.scr} = exp\ (1{,}620\ ln\ n_{lat} - 7{,}057)$
$n_{bg} = exp\ (1{,}901\ ln\ n_{lat} - 8{,}669)$
$n_{magh} = exp\ (1{,}914\ ln\ n_{lat} - 8{,}981)$
$n_{inc} = exp\ (2{,}000\ ln\ n_{lat} - 9{,}463)$
$n_{nec} = exp\ (2{,}404\ ln\ n_{lat} - 11{,}701)$
$n_{tc} = exp\ (2{,}513\ ln\ n_{lat} - 12{,}610)$
$n_{r(lat)} = exp(2{,}062\ ln\ n_{lat} - 10{,}713)$
$n_{gr} = exp\ (2{,}431\ ln\ n_{lat} - 12{,}625)$
$n_{onom} = exp(2{,}586\ ln\ n_{lat} - 13{,}794)$
$n_{scr} = exp\ (2{,}260\ ln\ n_{lat} - 12{,}266)$
$n_{germ} = exp\ (2{,}803\ ln\ n_{lat} - 15{,}722)$
$n_{it} = exp\ (2{,}986\ ln\ n_{lat} - 16{,}754)$
$n_{n.s.m} = exp\ (2{,}392\ ln\ n_{lat} - 13{,}713)$
$n_{rus} = exp\ (2{,}820\ ln\ n_{lat} - 16{,}233)$
$n_{ucr} = exp\ (3{,}580\ ln\ n_{lat} - 21{,}807)$



Scanned with OKEN Scanner

Fiindcă tot ne-am ocupat de problema fidelităţii modelului, să plasăm aici şi testul de adecvare căruia în cazul productivităţii lexicale îi rezervasem un paragraf aparte. Inventarierea numărului de sensuri ale cuvintelor cuprinse în DLRM ne-a condus la gruparea unităţilor lexicale în 31 de clase, dintre care primele 26 au fost prezentate exhaustiv în tabelul 16.1. Statistic vorbind, mai mult de jumătate dintre acestea (e vorba, evident, de cele care conţin cuvintele cele mai bogate în sensuri) sînt puţin semnificative din pricina sărăciei lor: clasele 1, 2, 3, 4, 5, 7, 8, 9, 12 şi 15 nu cuprind decît cîte un singur element, iar clasele 6, 10, 13, 14, 17 — cîte două. De altminteri originea unanim latină a termenilor din primele 17 clase face ca problema distribuţiei pe provenienţe să aibă, în domeniul respectiv, un caracter trivial. Începînd cu clasa a 18-a, numărul de cuvinte-bază, propriu şi cumulat, al claselor este cel prezentat în tabelul 17.2.

*Tabelul 17.2.*

| CLASA | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 6 | 6 | 9 | 15 | 14 | 27 | 37 | 72 | 113 | 229 | 454 | 1210 |
| Cumulat | 36 | 42 | 51 | 66 | 80 | 107 | 144 | 216 | 329 | 558 | 1012 | 2222 |

Repartiţia pe origini a cuvintelor din fiecare clasă ne este cunoscută din listele separate referitoare la fiecare aport lexical particular, pe care le-am prezentat in extenso în capitolul 16. Defalcate astfel pe provenienţe, cele 12 submulţimi al căror număr de elemente figurează în linia a treia a tabelului 17.2 devin tot atîtea segmente semnificative ale vocabularului ordonat după numărul de înţelesuri ale cuvintelor, ce pot servi drept valori de control al preciziei modelului $M_S$. Măsura în care deziderатul simulării matematice a ierarhiei induse de criteriul polisemiei în mulţimea cuvintelor româneşti este atins prin aplicarea ecuaţiilor elaborate îşi găseşte reflectarea în bilanţul comparativ ce face obiectul tabelului 17.3.

Parcurgerea două cîte două a liniilor tabelului evidenţiază un nivel acceptabil de concordanţă între valorile calculate şi cele reale. Îndeosebi în zonele unde intră în joc un număr mai mare de componente etimologice (către partea inferioară a tabelului), capacitatea modelului de a reda, cu o precizie absolut mulţumitoare, ponderea relativă a fiecărui aport lexical este manifestă. Ecarturile întîlnite mai cu seamă în prima parte a tabelului privesc comportamen-



Scanned with OKEN Scanner

*Tabelul 17.3*

| TOTAL CUVINTE | | LAT. | SLAV | cf. ALB. | „FR" | LAT. SAV. | BG. & SCR. | BG. | MAGH. | INC. | N.EC. | TC. | ROM. (LAT) | GR. | ONOM. | SCR. | GERM. | IT. | NSM | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 36 | calculat | 34 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 35 | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 42 | calculat | 38 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 40 | — | 1 | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 51 | calculat | 47 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 49 | — | 1 | — | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 66 | calculat | 59 | 2 | 1 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 60 | — | 1 | — | 3 | 1 | — | — | — | 1 | — | — | — | — | — | — | — | — | — | — |
| 80 | calculat | 71 | 3 | 1 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | real | 74 | — | 1 | — | 3 | 1 | — | — | — | 1 | — | — | — | — | — | — | — | — | — | — |
| 107 | calculat | 91 | 4 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| | real | 92 | 2 | 2 | 1 | 3 | 1 | — | 1 | — | 3 | — | 2 | — | — | — | — | — | — | — | — |
| 144 | calculat | 117 | 7 | 3 | 8 | 3 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| | real | 121 | 5 | 2 | 7 | 4 | 1 | — | 1 | — | — | — | 12 | — | 1 | — | — | — | — | — | — |
| 216 | calculat | 158 | 14 | 4 | 18 | 6 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — |
| | real | 161 | 12 | 3 | 19 | 4 | 1 | 1 | 3 | 2 | 5 | — | 2 | 1 | 2 | — | — | — | — | — | — |
| 329 | calculat | 219 | 24 | 6 | 39 | 11 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| | real | 221 | 23 | 3 | 42 | 7 | 1 | 2 | 6 | 4 | 9 | 1 | 3 | 2 | 2 | — | — | 1 | — | 1 | 1 |
| 558 | calculat | 317 | 48 | 10 | 92 | 22 | 10 | 10 | 8 | 8 | 9 | 6 | 3 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | — |
| | real | 323 | 49 | 7 | 93 | 21 | 4 | 9 | 9 | 8 | 15 | 4 | 3 | 4 | 4 | — | 1 | 1 | — | 2 | 1 |
| 1012 | calculat | 467 | 100 | 16 | 229 | 43 | 18 | 20 | 16 | 17 | 22 | 17 | 7 | 10 | 8 | 5 | 5 | 5 | 3 | 3 | 1 |
| | real | 465 | 113 | 16 | 218 | 52 | 17 | 21 | 16 | 17 | 23 | 11 | 7 | 10 | 8 | 2 | 4 | 5 | 2 | 4 | 1 |
| 2222 | calculat | 741 | 235 | 29 | 679 | 98 | 38 | 49 | 39 | 43 | 66 | 54 | 18 | 31 | 27 | 14 | 16 | 20 | 8 | 11 | 6 |
| | real | 753 | 240 | 29 | 675 | 93 | 43 | 45 | 40 | 44 | 64 | 56 | 19 | 32 | 28 | 8 | 17 | 18 | 7 | 8 | 3 |



Scanned with OKEN Scanner

tul cîtorva cuvinte cu statut mai aparte, a căror situaţie ambiguă a fost, în general, discutată anterior. Transferurile propuse, pentru un număr redus de termeni, de la fondul latin savant la cel „francez", de la cel bulgaro-sîrbocroat la cel vechi slav, de la cel cu etimologie obscură la cel autohton şi la cel slav rezolvă aproape în totalitate neconcordanţele locale ce mai apar şi care, oricum, nu ar putea umbri buna comportare de ansamblu a modelului.

## 18. Rezultatele aplicării modelului $M_S$. Paralela cu $M_p$

În intenţia de a face posibilă o comparaţie între rezultatele obţinute cu ajutorul modelelor $M_p$ şi $M_s$ pe întreaga întindere a vocabularului, ne vom raporta la aceleaşi repere numerice ca şi în cazul deja studiat al productivităţii lexicale. Presupunînd că vocabularul limbii române este o mulţime total ordonabilă pe criteriul numărului de sensuri ale cuvintelor, şi considerînd subşiruri descrescătoare iniţiale (altfel spus, fragmente neîntrerupte ale şirului care să înceapă toate cu primul său element) de lungime $n$, calculăm repartiţia după provenienţă a celor $n$ termeni ai subşirurilor respective pentru o gamă de variaţie a valorilor $n$ aleasă astfel:

— între $n =$ 0 şi $n =$ 100 din 10 în 10;
— între $n =$ 100 şi $n =$ 1 000 din 20 în 20;
— între $n =$ 1 000 şi $n =$ 2 000 din 50 în 50;
— între $n =$ 2 000 şi $n =$ 5 000 din 200 în 200;
— între $n =$ 5 000 şi $n =$ 10 000 din 1 000 în 1 000;
— între $n =$ 10 000 şi $n =$ 25 000 din 5 000 în 5 000

Paşii de calcul sînt, după cum se vede, identici cu cei din cazul modelului $M_p$. În tabelul 18.1 ordinea rubricilor verticale este ordinea intrării în scenă a diferitelor componente etimologice ale vocabularului. Astfel, primul element vechi slav apare la $n = 29$ cuvinte, primul termen de substrat — la $n = 32$, primul cuvînt de origine „franceză" — la $n = 38$ şi aşa mai departe. La fel procedasem şi în cazul clasificării după productivitate, dar ordinea obţinută acum diferă substanţial de cea stabilită atunci. Succesiunea apariţiilor, ce se



Scanned with OKEN Scanner

*Tabelul 18.1.*

| TOTAL | LAT. | SLAVE | cf.. ALB. | „FR" | LAT. SAV. | BG. & SCR. | BG. | MAGH. | INC. | NEC. | TC. | ROM. (LAT) | GR. | ONOM. | SCR. | GERM. | IT. | NSM | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 10 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 20 | 20 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 29 | 28 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 30 | 29 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 32 | 30 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 38 | 35 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 40 | 37 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 43 | 39 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 50 | 46 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 57 | 51 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 60 | 54 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 70 | 63 | 2 | 1 | 2 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 76 | 67 | 3 | 1 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 80 | 70 | 3 | 1 | 3 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 88 | 76 | 3 | 1 | 3 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 90 | 78 | 3 | 1 | 3 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 97 | 81 | 4 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 100 | 84 | 4 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 118 | 97 | 5 | 2 | 6 | 3 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 120 | 99 | 5 | 2 | 6 | 3 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 140 | 113 | 7 | 3 | 8 | 3 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 143 | 115 | 7 | 3 | 8 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 160 | 126 | 9 | 3 | 10 | 4 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **165** | 129 | 9 | 3 | 11 | 4 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — |
| 170 | 133 | 9 | 3 | 12 | 4 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — |
| **177** | 136 | 10 | 3 | 12 | 5 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — |
| 180 | 137 | 10 | 3 | 13 | 5 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — |
| 190 | 143 | 11 | 4 | 14 | 5 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — |
| 200 | 150 | 12 | 4 | 15 | 6 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — |
| **216** | 159 | 13 | 4 | 18 | 6 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — |
| 220 | 162 | 14 | 4 | 18 | 6 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — |
| **231** | 168 | 15 | 4 | 20 | 7 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 240 | 173 | 16 | 4 | 22 | 7 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 260 | 184 | 17 | 5 | 25 | 8 | 4 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 280 | 195 | 19 | 5 | 29 | 9 | 4 | 4 | 3 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| 300 | 205 | 22 | 6 | 33 | 10 | 5 | 4 | 3 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — |
| **320** | 215 | 23 | 6 | 36 | 11 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 2 | 1 | 1 | 1 | — | — | — | — |
| **325** | 217 | 24 | 6 | 37 | 11 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | | — | — |
| 340 | 223 | 25 | 6 | 40 | 12 | 6 | 5 | 4 | 4 | 4 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — |
| 357 | 232 | 27 | 7 | 44 | 12 | 6 | 5 | 4 | 4 | 4 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — |
| 360 | 233 | 28 | 7 | 45 | 12 | 6 | 5 | 4 | 4 | 4 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — |
| 380 | 242 | 30 | 7 | 48 | 13 | 6 | 6 | 5 | 5 | 5 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — |
| 393 | 248 | 31 | 7 | 51 | 14 | 7 | 6 | 5 | 5 | 5 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — |
| 400 | 251 | 32 | 7 | 53 | 14 | 7 | 6 | 5 | 5 | 5 | 4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — |
| 420 | 261 | 34 | 8 | 58 | 15 | 7 | 7 | 5 | 5 | 5 | 4 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — |
| 440 | 270 | 36 | 8 | 62 | 16 | 7 | 7 | 6 | 6 | 6 | 4 | 2 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | — |
| 460 | 278 | 38 | 8 | 67 | 17 | 8 | 8 | 6 | 6 | 6 | 5 | 2 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | — |
| 480 | 287 | 40 | 9 | 72 | 18 | 8 | 8 | 6 | 6 | 7 | 5 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | — |
| 500 | 295 | 42 | 9 | 77 | 19 | 9 | 9 | 7 | 7 | 7 | 5 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | — |



Scanned with OKEN Scanner

*Tabelul 18.1. (continuare)*

| TOTAL | LAT. | SLAVE | cf.. ALB. | „FR" | LAT. SAV. | BG. & SCR. | BG. | MAGH. | INC. | NEC. | TC. | ROM. (LAT) | GR. | ONOM. | SCR. | GERM. | IT. | NSM | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 520 | 303 | 45 | 9 | 82 | 20 | 9 | 9 | 7 | 7 | 7 | 6 | 3 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | — |
| 540 | 311 | 47 | 10 | 88 | 21 | 9 | 9 | 7 | 8 | 8 | 6 | 3 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | — |
| 560 | 318 | 49 | 10 | 92 | 22 | 10 | 10 | 8 | 8 | 9 | 6 | 3 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | — |
| 580 | 327 | 51 | 10 | 99 | 23 | 10 | 10 | 8 | 8 | 9 | 7 | 3 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | — |
| 600 | 334 | 53 | 10 | 103 | 24 | 11 | 11 | 9 | 9 | 10 | 7 | 4 | 4 | 3 | 2 | 2 | 2 | 1 | 1 | — |
| 620 | 341 | 55 | 11 | 109 | 25 | 11 | 11 | 9 | 9 | 10 | 8 | 4 | 5 | 4 | 2 | 2 | 2 | 1 | 1 | — |
| 640 | 349 | 58 | 11 | 114 | 26 | 11 | 12 | 9 | 9 | 11 | 8 | 4 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | — |
| 660 | 356 | 60 | 11 | 120 | 27 | 12 | 12 | 10 | 10 | 11 | 9 | 4 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | — |
| 680 | 363 | 62 | 12 | 126 | 28 | 12 | 13 | 10 | 10 | 12 | 9 | 4 | 6 | 4 | 3 | 2 | 2 | 1 | 1 | — |
| 686 | 365 | 63 | 12 | 128 | 28 | 12 | 13 | 10 | 10 | 12 | 9 | 4 | 6 | 4 | 3 | 2 | 2 | 1 | 1 | 1 |
| 700 | 370 | 65 | 12 | 132 | 28 | 12 | 13 | 10 | 11 | 12 | 9 | 4 | 6 | 4 | 3 | 2 | 2 | 2 | 2 | 1 |
| 720 | 376 | 66 | 12 | 136 | 29 | 13 | 14 | 11 | 11 | 13 | 10 | 5 | 6 | 5 | 3 | 2 | 3 | 2 | 2 | 1 |
| 740 | 383 | 69 | 12 | 144 | 30 | 13 | 14 | 11 | 11 | 13 | 10 | 5 | 6 | 5 | 3 | 3 | 3 | 2 | 2 | 1 |
| 760 | 389 | 71 | 13 | 149 | 31 | 14 | 14 | 11 | 12 | 14 | 11 | 5 | 7 | 5 | 3 | 3 | 3 | 2 | 2 | 1 |
| 780 | 396 | 73 | 13 | 156 | 32 | 14 | 15 | 12 | 12 | 15 | 11 | 5 | 7 | 5 | 3 | 3 | 3 | 2 | 2 | 1 |
| 800 | 403 | 76 | 13 | 161 | 33 | 14 | 15 | 12 | 13 | 15 | 12 | 5 | 7 | 6 | 4 | 3 | 3 | 2 | 2 | 1 |
| 820 | 409 | 78 | 14 | 167 | 34 | 15 | 16 | 13 | 13 | 16 | 12 | 5 | 7 | 6 | 4 | 3 | 3 | 2 | 2 | 1 |
| 840 | 416 | 80 | 14 | 174 | 35 | 15 | 16 | 13 | 13 | 16 | 13 | 6 | 8 | 6 | 4 | 3 | 3 | 2 | 2 | 1 |
| 860 | 422 | 82 | 14 | 181 | 36 | 15 | 17 | 13 | 14 | 17 | 13 | 6 | 8 | 6 | 4 | 3 | 4 | 2 | 2 | 1 |
| 880 | 428 | 84 | 14 | 186 | 37 | 16 | 17 | 14 | 14 | 18 | 14 | 6 | 8 | 7 | 4 | 4 | 4 | 2 | 2 | 1 |
| 900 | 434 | 87 | 15 | 192 | 38 | 16 | 18 | 14 | 15 | 18 | 14 | 6 | 9 | 7 | 4 | 4 | 4 | 2 | 2 | 1 |
| 920 | 440 | 89 | 15 | 199 | 39 | 17 | 18 | 14 | 15 | 19 | 15 | 6 | 9 | 7 | 4 | 4 | 4 | 2 | 3 | 1 |
| 940 | 446 | 92 | 15 | 206 | 40 | 17 | 19 | 15 | 15 | 19 | 15 | 6 | 9 | 7 | 5 | 4 | 4 | 2 | 3 | 1 |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 960 | 452 | 94 | 15 | 212 | 41 | 17 | 19 | 15 | 16 | 20 | 16 | 7 | 9 | 8 | 5 | 4 | 4 | 2 | 3 | 1 |
| 980 | 457 | 96 | 16 | 217 | 42 | 18 | 20 | 16 | 16 | 20 | 16 | 7 | 10 | 8 | 5 | 4 | 5 | 3 | 3 | 1 |
| 1000 | 464 | 98 | 16 | 225 | 42 | 18 | 20 | 16 | 17 | 21 | 17 | 7 | 10 | 8 | 5 | 4 | 5 | 3 | 3 | 1 |
| 1050 | 478 | 104 | 17 | 241 | 45 | 19 | 21 | 17 | 18 | 23 | 18 | 7 | 11 | 9 | 5 | 5 | 5 | 3 | 3 | 1 |
| 1100 | 492 | 110 | 17 | 258 | 47 | 20 | 23 | 18 | 19 | 24 | 19 | 8 | 12 | 9 | 6 | 5 | 6 | 3 | 3 | 1 |
| 1150 | 505 | 115 | 17 | 275 | 50 | 21 | 24 | 19 | 20 | 26 | 21 | 8 | 12 | 10 | 6 | 6 | 6 | 3 | 4 | 2 |
| 1200 | 519 | 121 | 18 | 293 | 52 | 21 | 25 | 20 | 21 | 27 | 22 | 9 | 13 | 11 | 6 | 6 | 7 | 3 | 4 | 2 |
| 1250 | 532 | 127 | 19 | 310 | 54 | 22 | 26 | 21 | 22 | 29 | 24 | 9 | 14 | 11 | 7 | 6 | 7 | 4 | 4 | 2 |
| 1300 | 544 | 132 | 19 | 327 | 57 | 23 | 27 | 22 | 23 | 31 | 25 | 10 | 15 | 12 | 7 | 7 | 8 | 4 | 5 | 2 |
| 1350 | 557 | 138 | 20 | 346 | 59 | 24 | 28 | 23 | 24 | 33 | 26 | 10 | 15 | 13 | 8 | 7 | 8 | 4 | 5 | 2 |
| 1400 | 569 | 144 | 20 | 363 | 61 | 25 | 30 | 24 | 25 | 34 | 28 | 11 | 16 | 14 | 8 | 8 | 9 | 4 | 5 | 2 |
| 1450 | 581 | 149 | 21 | 382 | 64 | 26 | 31 | 24 | 26 | 36 | 29 | 11 | 17 | 14 | 8 | 8 | 9 | 5 | 6 | 3 |
| 1500 | 592 | 155 | 21 | 399 | 66 | 27 | 32 | 25 | 27 | 38 | 31 | 12 | 18 | 15 | 9 | 9 | 10 | 5 | 6 | 3 |
| 1550 | 604 | 161 | 22 | 418 | 68 | 27 | 33 | 26 | 28 | 40 | 33 | 12 | 19 | 16 | 9 | 9 | 11 | 5 | 6 | 3 |
| 1600 | 615 | 166 | 23 | 437 | 70 | 28 | 34 | 27 | 29 | 42 | 34 | 13 | 20 | 17 | 9 | 10 | 11 | 5 | 7 | 3 |
| 1650 | 626 | 172 | 23 | 456 | 73 | 29 | 36 | 28 | 30 | 43 | 36 | 13 | 21 | 17 | 10 | 10 | 12 | 5 | 7 | 3 |
| 1700 | 637 | 177 | 24 | 474 | 75 | 30 | 37 | 29 | 32 | 45 | 37 | 13 | 22 | 18 | 10 | 11 | 13 | 5 | 7 | 4 |
| 1750 | 648 | 182 | 24 | 494 | 77 | 31 | 38 | 30 | 33 | 47 | 39 | 14 | 22 | 19 | 11 | 11 | 13 | 5 | 8 | 4 |
| 1800 | 658 | 188 | 24 | 513 | 80 | 32 | 39 | 31 | 34 | 49 | 40 | 14 | 23 | 20 | 11 | 12 | 14 | 6 | 8 | 4 |
| 1850 | 669 | 194 | 25 | 533 | 82 | 32 | 40 | 32 | 35 | 51 | 42 | 15 | 24 | 21 | 11 | 12 | 14 | 6 | 8 | 4 |
| 1900 | 678 | 199 | 26 | 551 | 84 | 33 | 41 | 33 | 36 | 53 | 44 | 15 | 25 | 22 | 12 | 13 | 15 | 6 | 9 | 5 |
| 1950 | 689 | 205 | 26 | 571 | 86 | 34 | 43 | 34 | 37 | 55 | 45 | 16 | 26 | 22 | 12 | 13 | 16 | 6 | 9 | 5 |
| 2000 | 698 | 210 | 27 | 591 | 88 | 35 | 44 | 35 | 38 | 57 | 47 | 16 | 27 | 23 | 13 | 14 | 16 | 7 | 9 | 5 |
| 2200 | 737 | 232 | 29 | 670 | 97 | 38 | 49 | 39 | 42 | 64 | 54 | 18 | 31 | 26 | 14 | 16 | 19 | 8 | 11 | 6 |
| 2400 | 774 | 255 | 30 | 751 | 106 | 41 | 53 | 42 | 46 | 72 | 60 | 20 | 35 | 30 | 16 | 19 | 22 | 9 | 12 | 7 |
| 2600 | 808 | 276 | 32 | 830 | 115 | 44 | 58 | 46 | 51 | 81 | 68 | 22 | 38 | 34 | 18 | 21 | 25 | 10 | 14 | 9 |



Scanned with OKEN Scanner

*Tabelul 18.1. (continuare)*

| TOTAL | LAT. | SLAVE | cf.. ALB. | „FR" | LAT. SAV. | BG. & SCR. | BG. | MAGH. | INC. | NEC. | TC. | ROM. (LAT) | GR. | ONOM. | SCR. | GERM. | IT. | NSM | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2800 | 841 | 297 | 34 | 915 | 123 | 47 | 62 | 50 | 55 | 90 | 75 | 24 | 42 | 37 | 19 | 23 | 29 | 11 | 16 | 10 |
| 3000 | 873 | 319 | 36 | 998 | 132 | 50 | 67 | 54 | 59 | 97 | 82 | 26 | 46 | 41 | 21 | 26 | 32 | 12 | 18 | 11 |
| 3200 | 904 | 339 | 37 | 1084 | 140 | 53 | 71 | 57 | 63 | 106 | 89 | 28 | 51 | 45 | 23 | 29 | 36 | 13 | 19 | 13 |
| 3400 | 933 | 360 | 39 | 1169 | 149 | 56 | 76 | 61 | 68 | 114 | 97 | 30 | 55 | 49 | 24 | 31 | 39 | 14 | 21 | 15 |
| 3600 | 962 | 382 | 41 | 1256 | 157 | 59 | 80 | 64 | 72 | 122 | 105 | 31 | 59 | 53 | 26 | 34 | 43 | 15 | 23 | 16 |
| 3800 | 990 | 402 | 42 | 1343 | 165 | 61 | 85 | 68 | 76 | 131 | 113 | 33 | 63 | 57 | 28 | 37 | 47 | 16 | 25 | 18 |
| 4000 | 1017 | 423 | 43 | 1435 | 173 | 64 | 89 | 71 | 80 | 139 | 120 | 35 | 67 | 61 | 29 | 40 | 50 | 17 | 27 | 20 |
| 4200 | 1042 | 442 | 45 | 1517 | 182 | 67 | 94 | 76 | 85 | 149 | 128 | 37 | 72 | 66 | 31 | 43 | 55 | 18 | 29 | 22 |
| 4400 | 1068 | 463 | 46 | 1606 | 189 | 70 | 99 | 79 | 89 | 158 | 136 | 39 | 76 | 70 | 33 | 46 | 59 | 19 | 31 | 24 |
| 4600 | 1092 | 483 | 48 | 1696 | 197 | 72 | 103 | 83 | 93 | 167 | 144 | 41 | 81 | 74 | 35 | 49 | 63 | 20 | 33 | 26 |
| 4800 | 1117 | 504 | 49 | 1790 | 205 | 74 | 107 | 86 | 97 | 175 | 153 | 43 | 84 | 78 | 36 | 52 | 67 | 21 | 35 | 27 |
| 5000 | 1141 | 524 | 50 | 1879 | 213 | 77 | 111 | 89 | 101 | 185 | 161 | 45 | 89 | 82 | 38 | 55 | 71 | 22 | 37 | 30 |
| 6000 | 1241 | 622 | 56 | 2339 | 251 | 90 | 133 | 106 | 121 | 231 | 202 | 54 | 111 | 104 | 47 | 71 | 94 | 28 | 48 | 41 |
| 7000 | 1351 | 718 | 62 | 2804 | 288 | 101 | 153 | 123 | 142 | 278 | 246 | 63 | 134 | 127 | 56 | 88 | 118 | 34 | 60 | 54 |
| 8000 | 1442 | 810 | 68 | 3271 | 324 | 113 | 174 | 140 | 162 | 326 | 290 | 73 | 158 | 151 | 65 | 107 | 144 | 40 | 73 | 69 |
| 9000 | 1528 | 903 | 73 | 3750 | 359 | 124 | 194 | 157 | 182 | 372 | 335 | 82 | 181 | 175 | 74 | 125 | 170 | 46 | 85 | 85 |
| 10000 | 1608 | 992 | 82 | 4228 | 394 | 135 | 214 | 173 | 201 | 420 | 380 | 91 | 205 | 200 | 83 | 145 | 198 | 51 | 98 | 102 |
| 15000 | 1951 | 1421 | 100 | 6679 | 557 | 184 | 309 | 249 | 295 | 663 | 619 | 135 | 327 | 329 | 128 | 248 | 353 | 81 | 169 | 203 |
| 20000 | 2231 | 1824 | 118 | 9164 | 708 | 229 | 398 | 322 | 386 | 907 | 868 | 178 | 454 | 466 | 174 | 361 | 527 | 110 | 247 | 328 |
| 25000 | 2470 | 2205 | 136 | 11648 | 850 | 270 | 484 | 392 | 474 | 1181 | 1120 | 220 | 582 | 607 | 219 | 481 | 715 | 144 | 329 | 473 |



Scanned with OKEN Scanner

constituie de fapt într-o ierarhie de importanţă a aporturilor lexicale alogene, este prezentată în paralel pentru cele două modele în tabelul 18.2

*Tabelul 18.2*

| $n$ | 5 | 9 | 29 | 31 | 32 | 38 | 39 | 40 | 43 | 54 | 57 | 76 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $M_p$ | cf. alb. | slav | — | inc. | — | — | onom. | bg. & scr. | — | bg. | magh. | — |
| $M_s$ | | — | slav | — | cf. alb. | „fr" | — | — | lat. sav. | — | bg. & scr. | bg. |

| $n$ | 78 | 88 | 97 | 117 | 118 | 135 | 143 | 165 | 177 | 216 | 231 | 254 | 276 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $M_p$ | nec. | — | — | nsm | — | gr. | — | — | — | — | — | germ. | tc. |
| $Ms$ | — | magh. | inc. | — | nec. | — | tc. | rom (lat) | gr. | onom. | scr. | — | — |

| $n$ | 320 | 325 | 347 | 353 | 357 | 384 | 393 | 405 | 546 | 585 | 613 | 686 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Mp$ | — | — | scr. | lat. sav. | — | ucr. | — | „fr" | rom. (lat.) | rus | it. | — |
| $M_s$ | germ. | it. | — | — | nsm | — | rus | — | — | — | — | ucr. |

Ce observaţii se pot face pe marginea acestei duble înşiruiri?

O primă remarcă, cu caracter mai general, priveşte ambitusul zonei de inserţie a aporturilor respective, care este, în mod evident, mai restrîns în cazul al doilea decît în cel dintîi. În clasificarea după gradul de polisemie, primul element non-latin apare mai tîrziu decît în clasificarea după numărul de derivate şi, cu toate acestea, practic toate componentele etimologice ale limbii române (cu excepţia nesemnificativă a elementului ucrainian) îşi fac apariţia înainte de rangul $n = 400$, în timp ce pe lista productivităţii, punctele de ocurenţă sînt mult mai dispersate: substratul urcă pînă la rangul $n = 5$, dar primul împrumut din italiană se întîlneşte de-abia la $n = 613$. Această împrăştiere mai mare trebuie, fără doar şi poate, pusă în legătură cu dependenţa de vechime a numărului de derivate, condiţionare absentă în cazul numărului de sensuri, din motivul, deja arătat, că adesea cuvintele împrumutate pătrund în limbă împreună cu întregul buchet



Scanned with OKEN Scanner

de sensuri pe care le dobîndiseră în limba de origine. Dacă un neologism recent este întotdeauna plasat foarte jos în ierarhia productivităţii, el poate ocupa, în schimb, un loc bun în cea a polisemiei, totul depinzînd de cîte înţelesuri ajunsese să posede etimonul său în idiomul-sursă. Aşa se explică importantele răsturnări de situaţii din tabelul 18.2, dintre care cea mai izbitoare este ascensiunea elementului „francez" de la rangul $n = 405$ la $n = 38$. Materialul neologic grupat în această categorie se dovedeşte astfel cel mai însemnat ca pondere semantică după fondul primordial şi aportul vechi slav.

Din punctul de vedere al poziţiei relative ocupate de primul lor reprezentant în cele două clasamente considerate, componentele etimologice ale vocabularului pot fi grupate în trei clase:

a. Contribuţii cu poziţie identică în ambele ierarhii. Din această categorie fac parte aportul bulgaro-sîrbocroat, care intră în scenă al 5-lea în ambele clasificări, cel bulgăresc (al 6-lea), cel maghiar (al 7-lea) şi cel rusesc (al 18-lea). Putem asimila aceleiaşi clase, fără a greşi prea mult, substratul traco-dacic şi elementul vechi slav, prea apropiate ca poziţie pentru ca rocada lor să poată fi privită drept semnificativă.

b. Contribuţii care înregistrează o creştere (de regulă spectaculoasă) în ierarhia polisemiei faţă de poziţia pe care o ocupau în clasificarea după productivitate. Remarcabil este faptul că aproape toate aporturile aflate în această situaţie sînt de sorginte latină. E vorba de elementul „francez" (care trece de pe locul 16 pe locul 3!), de cel latin savant (de pe 14 pe 4), de cel italian (de pe 19 pe 16) şi de formaţiile româneşti din material lexical latin (de pe 17 pe 11). Ele contribuie în mod esenţial la consolidarea romanităţii vocabularului. Li se mai adaugă elementul turcesc, ce nu cîştigă însă decît două locuri în clasament (de la 12 la 10).

c. Contribuţii care suferă o anumită depreciere, retrogradînd pe poziţii mai modeste. Astfel elementul sîrbocroat şi cuvintele cu etimologie obscură pierd cîte un loc, trecînd de pe poziţia a 13-a pe cea de a 14-a şi, respectiv, de pe poziţia a 8-a pe a 9-a, elementul grecesc — două locuri (de pe 10 pe 12), cel ucrainian — patru locuri (de pe 15 pe 19), materialul lexical de origine incertă — cinci locuri (de pe 3 pe 8), cel zis „nord-slav medieval" — opt locuri (de pe 9 pe 17). Căderea cea mai impresionantă o cunosc formaţiunile onomatopeice care pierd 9 locuri (de la 4 la 13), fenomen uşor de interpretat, avînd în vedere, pe de o parte, că ele sînt în majoritate vechi şi deci au avut timp să producă un număr mare de derivate şi,

pe de alta, că presiunea formei interne datorate caracterului lor imitativ îngrădeşte drastic posibilităţile de generare a unor înţelesuri noi, divergente.

Multe dintre caracteristicile şi tendinţele generale relevate de modelul $M_p$ se menţin şi în cazul modelului $M_s$. Astfel, supremaţia fondului latin moştenit în zona nucleară a vocabularului reprezentată de primele linii ale tabelului 18.1 este încă mai izbitoare decît rezultase din studiul productivităţii lexicale. Punerea faţă în faţă (tabelul 18.3), pentru valori $n$ progresiv crescătoare, a cifrelor absolute reprezentînd componenta lexicală latină ilustrează convingător această afirmaţie.

*Tabelul 18.3*

| $n$ | | | 10 | 20 | 30 | 40 | 50 | 100 |
|---|---|---|---|---|---|---|---|---|
| Nr. de cuv.-bază de origine latină | $M_p$ | absolut | 8 | 17 | 25 | 31 | 38 | 72 |
| | | % | 80,0 | 85,0 | 83,3 | 77,5 | 76,0 | 72,0 |
| | $M_s$ | absolut | 10 | 20 | 29 | 37 | 46 | 84 |
| | | % | 100,0 | 100,0 | 96,7 | 92,5 | 92,0 | 84,0 |

| $n$ | | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|
| $M_p$ | absolut | 135 | 191 | 242 | 292 | 336 | 381 | 421 | 460 | 498 |
| | % | 67,5 | 63,7 | 60,5 | 58,4 | 56,0 | 54,4 | 52,6 | 51,1 | 49,8 |
| $M_s$ | absolut | 150 | 205 | 251 | 295 | 334 | 370 | 403 | 434 | 464 |
| | % | 75,0 | 68,3 | 62,8 | 59,0 | 55,7 | 52,9 | 50,4 | 48,2 | 46,4 |

Se observă că, pînă în jurul rangului $n = 550$, predominanţa elementului latin este mai pregnantă în clasificarea după gradul de polisemie decît în cea după productivitate. Datorat în principal prezenţei unor categorii morfologice inapte de a genera derivate (pronume, articole, conjuncţii, prepoziţii), acest decalaj se atenuează treptat, ajungîndu-se ca momentul cînd fondul latin îşi pierde majoritatea absolută să survină mai devreme pentru $M_s$ (pe la $n = 810$ decît pentru $M_p$ (în jur de $n = 990$).

Se observă că, pînă în jurul rangului $n = 550$, predominanţa elementului latin este mai puternică în clasificarea după gradul de polisemie decît în cea după productivitate.

Să urmărim însă ce se întîmplă şi la valori superioare ale parametrului $n$ (tabelul 18.4).



Scanned with OKEN Scanner

*Tabelul 18.4*

| $n$ | | | 1000 | 2000 | 3000 | 4000 | 5000 | 6000 |
|---|---|---|---|---|---|---|---|---|
| Nr. de cuv.-bază de origine latină | $M_p$ | absolut | 498 | 792 | 992 | 1139 | 1257 | 1353 |
| | | % | 49,8 | 39,6 | 33,1 | 28,5 | 25,1 | 22,6 |
| | $M_s$ | absolut | 464 | 698 | 873 | 1017 | 1141 | 1251 |
| | | % | 46,4 | 34,9 | 29,1 | 25,4 | 22,8 | 20,9 |

| $n$ | | 7000 | 8000 | 9000 | 10000 | 15000 | 20000 | 25000 |
|---|---|---|---|---|---|---|---|---|
| $M_p$ | absolut | 1435 | 1505 | 1569 | 1625 | 1848 | 2009 | 2138 |
| | % | 20,5 | 18,8 | 17,4 | 16,3 | 12,3 | 10,0 | 8,6 |
| $M_s$ | absolut | 1351 | 1442 | 1528 | 1608 | 1951 | 2231 | 2470 |
| | % | 19,3 | 18,0 | 17,0 | 16,1 | 13,0 | 11,1 | 9,9 |

Tendinţa constatată anterior, de reducere mai rapidă a ponderii elementului latin în clasamentul polisemiei decît în cel al productivităţii, nu se prelungeşte indefinit. Dimpotrivă. Există un prag (cu puţin peste $n = 10\,000$) unde cotele părţi reprezentate de termenii moşteniţi din latineşte devin din nou egale, după care modelul productivităţii prezintă din nou fondul latin al vocabularului într-o lumină mai favorabilă decît modelul polisemiei. Pe seama cărui factor să punem această nouă răsturnare de situaţie?

Orice încercare de explicare a unor atari schimbări trebuie să aibă în vedere faptul că dinamica distribuţiei procentuale a celor 12 componente etimologice principale ale limbii depinde simultan de ritmurile de creştere ale tuturor acestor componente. Însă chiar dacă toate au, în principiu, un cuvînt de spus, nu este mai puţin adevărat că faciesul general al vocabularului e determinat cu precădere de aporturile cele mai dinamice, mai active. Or, componenta care prezintă în cel mai înalt grad această calitate este cea „franceză". Debutînd la valori $n$ mult mai mici în modelul $M_s$ decît în modelul $M_p$, ea intră repede în concurenţă cu elementul latin, a cărui pondere ajunge să o egaleze în jurul rangului $n = 2\,500$, ceea ce nu se întîmplă în ierarhia productivităţii decît la ceva mai mult de $n = 5\,000$. Reducerea timpurie la un rol secundar a iniţial atît de puternicului factor latin se datorează tocmai acestei ascensiuni impetuoase. Pe de altă parte, producţia modestă de derivate, specifică neologismelor, menţine pentru o bună bucată de vreme elementul „francez" într-o poziţie subordonată pe

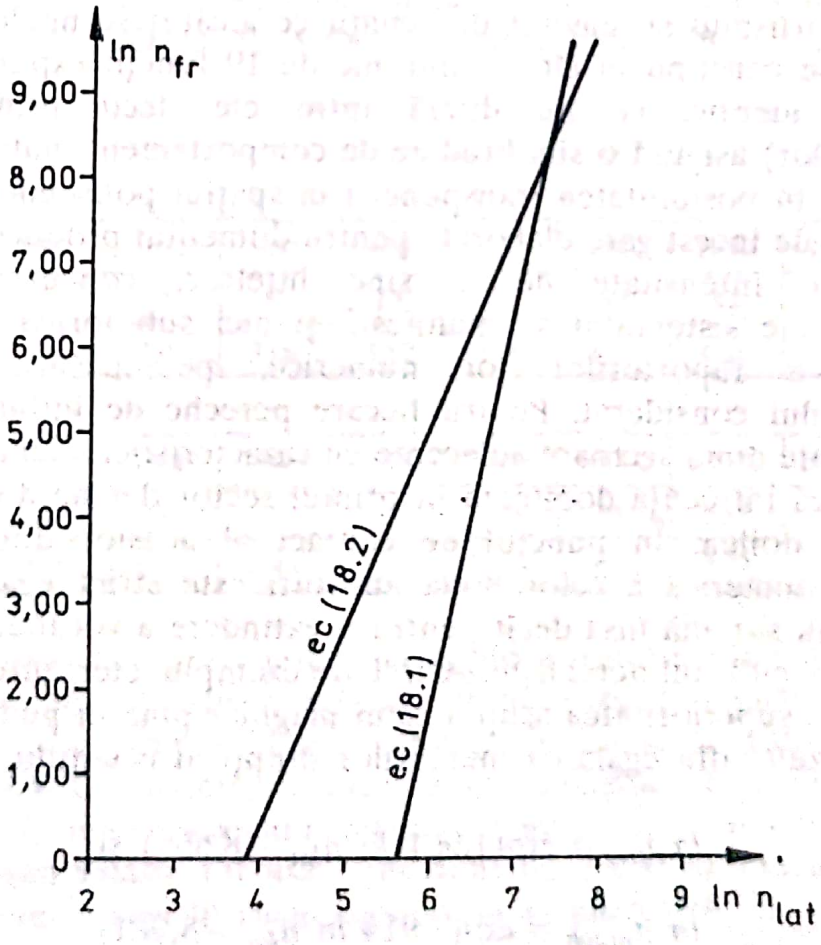
scara productivităţii lexicale, perioadă urmată de o adevărată „explozie demografică", a cărei explicaţie ne-o oferă compararea între ele a ecuaţiilor, aferente domeniului „francez", ce intră în alcătuirea modelelor $M_p$ şi $M_s$. Să ni le reamintim în forma lor logaritmată:

$$M_p : \ln n_{fr} = 4{,}748 \ln n_{lat} - 26{,}786 \qquad (18.1)$$

$$M_s : \ln n_{fr} = 2{,}361 \ln n_{lat} - 9{,}081 \qquad (18.2)$$

Reprezentate pe o aceeaşi diagramă, (fig. 18.1), dreptele ce le corespund sînt concurente în punctul de coordonate (7,417 ; 8,431). Măsura dinamismului componentei „franceze" este dată de cele două pante, $m_1 = 4{,}748$ şi $m_2 = 2{,}361$, prima mult superioară celei de a doua, ceea ce explică expansiunea imediat următoare momentului de echilibru care corespunde aproximativ valorii $n = 10\,000$. Evoluţia procentului reprezentat de fondul latin este strict dependentă de această variaţie, deoarece, după cum s-a arătat, începînd încă de la $n = 2\,500$, pentru clasificarea după numărul de derivate, şi $n = 5\,000$, pentru cea după numărul de sensuri, materialul neologic şi-a asigurat

FIG. 18.1



preponderenţa asupra celui moştenit şi orice creştere procentuală a primului se repercutează corespunzător asupra celui de la doilea.

Simplificată la maximum prin raportarea la numai două componente etimologice ale vocabularului, cea latină moştenită şi cea „franceză", în intenţia de a face inteligibile tendinţele destul de complexe ce se manifestă în sînul unui vocabular gîndit ca extensibil, pe măsura lărgirii graniţelor sale, analiza de mai sus poate fi continuată, luîndu-se în considerare pe rînd sau simultan şi variaţia altor aporturi lexicale, întreprindere migăloasă şi, fireşte, consumatoare de spaţiu editorial, la care renunţăm şi pe considerentul că, oricum, cititorul dispune în prezent de toate datele pentru a efectua singur toate comparaţiile ce îl interesează.

Să reţinem numai că şi perspectiva semantică plasează elementul autohton şi pe cel vechi slav în apropierea nemijlocită a fundamentului latin al vocabularului şi că, spre deosebire de studiul productivităţii lexicale, ea acordă un rol de prim plan influenţei neologice de sorginte direct sau mediat latină (termeni provenind din latina literară, din franceză şi din italiană).

Omomorfismul sistemelor de ecuaţii ce alcătuiesc modelele $M_p$ şi $M_s$ (ambele constînd în cîte o mulţime de 19 funcţii exponenţiale de argument identic, ce nu diferă între ele decît prin valoarea coeficienţilor) asigură o similitudine de comportament matematic care se traduce în posibilitatea transpunerii în spaţiul polisemiei a tuturor metodelor de investigare elaborate pentru domeniul productivităţii. Cu variaţii de intensitate de la sine înţelese, concurenţa dintre componentele sistemului se manifestă şi aici sub forma modificării continue a raporturilor lor numerice, pe măsura extinderii vocabularului considerat. Pentru fiecare pereche de influenţe, există cu necesitate două sectoare adiacente cu caracteristici diametral opuse, în sensul că influenţa dominată în primul sector devine dominantă în cel de al doilea. În punctul de contact al acestor două domenii ponderea numerică a celor două aporturi este strict egală, ceea ce poate să nu survină însă decît pentru o extindere a vocabularului pînă la proporţii cu totul nerealiste. Astfel, de exemplu, elementul bulgăresc îşi menţine superioritatea asupra celui maghiar pînă în punctul a cărui abscisă rezultă din egalarea membrilor drepţi ai ecuaţiilor:

$$ln\ n_{bg} = exp(1{,}901\ ln\ n_{lat} - 8{,}669)\ \text{şi}$$

$$ln\ n_{magh} = exp(1{,}914\ ln\ n_{lat} - 8{,}981)$$



Scanned with OKEN Scanner

Rezultă de îndată $ln\ n_{lat} = 24$, de unde concluzia că aportul maghiar ajunge egal cu cel bulgăresc de-abia într-un vocabular cuprinzînd cca. 25.600.000.000 de cuvinte de origine latină. Menţionăm, ca o curiozitate, că într-un astfel de lexic ipotetic, elementul dominant l-ar reprezenta cuvintele de obîrşie ucraineană, al căror număr s-ar cifra la 6.561.000.000.000.000.000.000.000.000, pe locul al doilea situîndu-se termenii proveniţi din italieneşte! O română ruteano-peninsulară — iată ciudăţenia la care ne poate conduce extrapolarea dincolo de orice limită rezonabilă a tendinţelor absolut reale exprimate de „paradoxul lui A. de Cihac".

Un alt corolar al modelului $M_p$ pe care îl regăsim şi în cazul omologului său $M_s$ se referă la existenţa pentru fiecare componentă etimologică a vocabularului a unui punct de maximum procentual, adică a unei anverguri strict determinate a vocabularului în care cota parte reprezentată de componenta respectivă îşi atinge plafonul maxim. Pentru ilustrare, să considerăm, de data aceasta, evoluţia ponderii elementului latin savant. Extrăgînd selectiv din tabelul 18.1 cîteva valori caracteristice ale acestei contribuţii pentru o gamă largă de variaţie a lui $n$ ($n = 50 - 25\,000$) şi calculînd de fiecare dată procentul reprezentat de acestea din volumul lexical total, obţinem situaţia din tabelul 18.5.

*Tabelul 18.5.*

| $n$ | 50 | 100 | 200 | 400 | 600 | 800 | 1000 | 1100 | 1200 |
|---|---|---|---|---|---|---|---|---|---|
| $n_{lat.sav.}$ | 1 | 2 | 6 | 14 | 24 | 33 | 42 | 47 | 52 |
| % | 2,00 | 2,00 | 3,00 | 3,50 | 4,00 | 4,13 | 4,20 | 4,27 | 4,33 |

| $n$ | 1400 | 1600 | 1800 | 2000 | 5000 | 10000 | 15000 | 20000 | 25000 |
|---|---|---|---|---|---|---|---|---|---|
| $n_{lat.sav.}$ | 61 | 70 | 80 | 88 | 213 | 394 | 557 | 708 | 850 |
| % | 4,36 | 4,38 | 4,44 | 4,40 | 4,26 | 3,94 | 3,71 | 3,54 | 3,40 |

Se observă o creştere monotonă a procentului neologismelor latine pînă în dreptul valorii $n = 1\,800$, după care însemnătatea acestei participări scade treptat. Un vocabular limitat la cele mai bogate în sensuri 1 800 de cuvinte flatează aşadar imaginea contribuţiei latine literare la formarea lexicului românesc modern, punînd-o în lumina cea mai avantajoasă, tot aşa cum unul de 1 100 de cuvinte favoriza elementul vechi slav în cazul modelului productivităţii.



Scanned with OKEN Scanner

## 19. Numărul de expresii şi locuţiuni ca indicator al importanţei cuvintelor

Ajunşi la acest al treilea criteriu de clasificare a cuvintelor, este cazul să ne întrebăm dacă el constituie sau nu un parametru independent de cei consideraţi pînă acum (productivitatea şi polisemia).

Răspunsul, uşor de dat, e categoric negativ. Participarea la expresii se află în mod evident într-o strînsă intercondiţionare cu polisemia. Deschizînd orice dicţionar explicativ, avem prilejul să constatăm că printre sensurile socotite drept de sine stătătoare ale unui cuvînt o bună parte o reprezintă accepţiunile pe care acesta le capătă exclusiv în contextele fixe determinate de anumite unităţi frazeologice sau locuţiuni.

Astfel, dintre cele 5 înţelesuri ale substantivului *mînă* menţionate în DLRM, două (1. „fiecare dintre cele două membre superioare ale omului, de la umăr pînă la vîrful degetelor, în special partea de la extremitatea de jos a antebraţului"; 2. „cantitate din ceva, cît poate fi cuprinsă în palmă") au putut fi definite în afara oricărui context, unul („tărie fizică, forţă, putere") se descoperă numai în unele expresii („mînă de lucru", „a ajunge pe mîini bune" „a avea pe cineva sub mînă", „a încăpea pe mîna cuiva", „a avea mînă de fier"), altul într-o singură unitate frazeologică („mînă curentă") şi, în fine, cel din urmă într-un compus („mîna-Maicii-Domnului"). Existenţa expresiilor poate deci spori polisemia unui cuvînt, dar legătura este departe de a fi simplă, căci un termen poate figura în multe expresii fără a-şi schimba sensul de bază, după cum poate să prezinte o multitudine de înţelesuri necondiţionate frazeologic (tipurile pure sînt totuşi greu de exemplificat, dovadă că şi această afirmaţie ar trebui nuanţată). Să menţionăm, în această ordine de idei, cuvîntul monosemantic *şarpe,* prezent într-un mare număr de expresii şi locuţiuni („ca din gură de şarpe", „(a sări) ca muşcat de şarpe", „a fuma ca un şarpe", „gol ca şarpele", „a încălzi şarpele la sîn", „limbă de şarpe", „a-l muşca (pe cineva) şarpele de inimă", „pui de şarpe" etc) şi, la celălalt pol, termenul latin literar *coroană* pentru care dicţionarul indică 8 sensuri diferite, dar nici o expresie.

S-ar părea deci că între polisemie şi participarea la expresii subzistă atît o anumită suprapunere cît şi un oarecare grad de complemen-



Scanned with OKEN Scanner

taritate. În orice caz, inventarierea expresiilor şi a locuţiunilor contribuie la mai buna situare a cuvintelor în ierarhia ideală spre care năzuim. Este motivul pentru care ne ocupăm şi de acest aspect, încercînd să îl tratăm, în ciuda caracterului incomplet al datelor de bază[1], conform aceleiaşi scheme metodologice pe care am utilizat-o în studiul productivităţii şi al polisemiei.

Pentru a oferi mai întîi o idee despre raportul global dintre elementul latin moştenit şi aporturile ulterioare, iată mai jos lista ordonată descrescător a cuvintelor care participă în DEL la cel puţin 4 expresii. Caracterele cursive evidenţiază contribuţia non-latină care creşte progresiv cu rangul cuvintelor.

| | *Unităţi lexicale.* | *Nr. expr. (DEL)* |
|---|---|---|
| 1. | ţine | 84 |
| 2. | ochi | 82 |
| 3. | da (vb.) | 71 |
| 4. | cap | 69 |
| 5. | lua | 65 |
| 6. | mînă | 63 |
| 7. | gură | 58 |
| 8—9. | avea, pune | 52 |
| 10—11. | nas, picior | 51 |
| 12. | ridica | 48 |
| 13—14. | bate, face | 46 |
| 15. | *vorbă* | 45 |
| 16. | pămînt | 44 |
| 17. | zi | 42 |
| 18. | trage | 39 |
| 19. | sta | 38 |
| 20—21. | inimă, loc | 35 |
| 22. | scoate | 34 |
| 23—24. | drac, rău | 32 |
| 25. | bine | 31 |
| 26—27. | bun, veni | 30 |
| 28—29. | parte, *rînd* | 29 |

[1] Un dicţionar exhaustiv de expresii româneşti lipseşte, din păcate, deocamdată. În ceea ce ne priveşte, folosim aici *Dicţionarul de expresii şi locuţiuni ale limbii române* (DEL) de Gabriela Duda, Aglaia Gugui şi Marie Jeanne Wojcicki, Editura Albatros, 1985.



Scanned with OKEN Scanner

30. sînge . . . . . 28
31. suflet . . . . . 27
32—35. cuvînt, duce, *drum,* mînca . . . . . 26
36—38. cădea, drept, *nevoie* . . . . . 25
39—41. apă, *băga, vreme* . . . . . 24
42—49. aşa, faţă, intra, limbă, minte, *obraz, seamă,* şti . 23
50—51. *gînd,* trece . . . . . 22
52—54. foc, rămîne, ureche . . . . . 21
55—57. piele, prinde, sus . . . . . 20
58—61. cale, greu, măsură, mult . . . . . 19
62—67. deschide, lume, merge, păr, urmă, *voie* . . . . . 18
68—69. lăsa, ţară . . . . . 17
70—76. atît, coadă, deget, întinde, întoarce, nume, strînge 16
77—87. fi, jos, moară, moarte, mort, oală, punct, şedea, tăia, tot, zice . . . . . 15
88—100. an, călca, *chip,* cruce, *gît,* ieşi, lumină, timp, *treabă,* unu, uşă, viaţă, vînt . . . . . 14
101—114. alb, cît, *drag,* lega, mijloc, muscă, os, pace, palmă, păcat, piatră, piept, sări, vedere . . . . . 13
115—125. cer, doi, înainte, *milă,* negru, pas, pierde, *pica,* rupe, scăpa, şarpe. . . . . 12
126—140. ajunge, apuca, cal, colo, dinte, dos, *fel, găsi, necaz, noroc,* pîine, roată, *socoteală,* vale, *zor* . . . . . 11
141—160. altul, arunca, azi, *ban,* cum, fugă, gust, înghiţi, lup, mare (adj.), om, *pomană,* puţin, rade, rece, rost, sare, spate, *tobă,* vărsa . . . . . 10
161—184. aduce, amar, capăt, coastă, Dumnezeu, fire, *gol,* larg, lung, măsea, oaie, *opincă,* ou, *pînză, popă,* preţ, *sfînt,* sparge, strica, sui, şapte, şterge, umbla, uşor . . . 9
185—211. arde, bătaie, *bot,* ce, *ceas, cojoc,* corn, crede, dată, *glumă,* lege, *nădejde, neam,* pai, patru, părere, *păzi, pisică, poveste,* putere, rîs, scurt, soare, umbră, *veac, vîrî,* vrea . . . . . 8
212—256. aer, aur, *băţ, bob, buză, căciulă,* căuta, cînd, crăpa, *cumpănă,* furcă, *gheară, glas,* gros, *haz,* încurca, înţelege, joc, juca, lapte, lat, lemn, lucru, lună (astr.), mînecă, *mîţă,* noapte, număra, oară, *perdea, poftă,* putea, *răboj,* suflare, sîn, *talpă,* tare, tată, *trăi,* umăr, unul, urît, *val, vină,* zeamă . . . . . 7
257—322. ac, afla, armă, aţă, bucată, carne, carte, cămaşă, *ciolan,* cîine, *colac,* cot, cunoştinţă, curat, departe,



Scanned with OKEN Scanner

Domn, fiinţă, fugi, *gard, grai, grijă, idee,* închide, *jumătate,* lacrimă, *leac,* miere, minune, muia, muri, mustaţă, nod, *notă,* nouă, nu, omenie, *omorî, onoare, oră,* orb, *papară, para, petic, plesni, poală,* poamă, porc, *praf, praştie,* purta, putinţă, rîde, *ruşine,* scăpăra, schimba, scurta, *silă, slab,* stea, tăcea, toacă, tunde, unghie, *vai,* vedea, *vorbi* 6

323—416. acela, aici, arăta, ăsta, *ba, balama, bălai,* bou, *brînză, burduf, buzunar,* car, care, călcîi, căpătîi, cărare, ceva, *chef, cinste, circulaţie,* coardă, *cont,* creşte, cruciş, curs, cuţit, drege, după, fătă, floare, frînge, funie, înşira, învîrti, jur, latură, mamă, masă, mic, mormînt, *moş,* naştere, *neamţ,* nimic, ninge, număr, *olaltă, pagubă,* pană, *parolă,* paşte, părea, peşte, *pic, picătură, piper,* pîntece, plin, *prag, raport, rezervă,* roade, roşu, *rudă,* semn, *slobod,* spaimă, spăla, sprînceană, spune, strîmb, subţire, suna, *sută, şagă,* şoarece, *ştiinţă, şurub, temei, teren, termen, tîrg, toană, trai, traistă,* trei, tun, turtă, unde, unge, *vatră, veste,* vînă, *vîrf* . . . . . 5

417—560. *act,* afară, aminte, aproape, *astăzi,* astupa, aşterne, *atenţie,* auzi, ăla, baieră, *balanţă,* barbă, *bază,* bea, *boală, bogdaproste, borş, brazdă,* brînci, bucăţică, *buche, butuc,* casă, căruţă, *ceafă,* cenuşă, cere, chema, *chior, ciudă,* cine, cînta, coace, *cobză, condei, coş, credit,* cuc, cunoaşte, curge, *curmeziş, darabană, deal,* doagă, *dragoste, duh, Dunăre, egal,* falcă, fier, *forţă, frică,* frîu, fund, găină, genunchi, gheaţă, *groapă, grumaz, haram,* încă, începea, încoace, încredere, îndoială, întreg, înţeles, *liber,* lungi, *materie,* măsura, minciună, mira, moale, muchie, muşca, *mutră, naiba,* naşte, *natură, năvală, nerv, nevoinţă,* numai, nuntă, *olac, ordin, ordine, pahar, papuc,* par, păţi, perete, *pită,* pleca, ploua, *pod, ponos, poştă, proces,* pumn, purice, *pustiu, răbdare,* răsări, *respect, reteza, rînză,* scrie, *scump,* sec, *serviciu, sfoară,* simţi, sorţi, stare, *stăpîn, steag,* stîng, strîmba, *strună,* sufla, suge, surd, *şatră,* ştire, *taină, taler,* tărie, *tipar,* toarce, turna, *ţigan, ţintă,* umfla, , umple, urs, venin, verde, *vileag,* viu, *zare,* zbura . . . . . 4

Lista de mai sus este alcătuită precumpănitor din termeni-matcă, dar în zona rangurilor superioare îşi fac treptat apariţia şi unele cuvinte derivate. Ca şi în cazul clasificării după numărul de sensuri, cele dintîi monopolizează poziţiile fruntaşe: primele 100 de unităţi sînt, fără excepţie, cuvinte-bază: de-abia începînd cu rangul 101 se ivesc, răzleţ, şi termenii derivaţi. Pînă la nivelul cuvintelor ce participă la cîte 4 expresii diferite, ponderea acestora din urmă rămîne modestă. Avînd în vedere acest fapt, dar şi, mai ales, intenţia mărturisită de a compara în final între ele ierarhiile induse de diferitele criterii considerate, ne mărginim şi de data aceasta la analiza situaţiei cuvintelor nederivate.

Pe cele de origine latină nu le mai enumerăm separat, deoarece le-am delimitat deja de restul cuvintelor prin mijlocirea unor caractere tipografice deosebitoare. Rămîne să le prezentăm numai pe acelea care intră în alcătuirea a mai puţin de 4 expresii şi locuţiuni. Iată-le (tot pe baza materialului din DEL):

344—439. acesta, ajuta, albastru, alege, apropia, arici, atunci, băutură, biserică, braţ, broască, cald, cel, cheag, cheie, chingă, ciur, cîntec, cui, culca, cumpăt, cuviinţă, da, dezlega, dimineaţă, dormi, dulce, durea, el, făină, frige, frumos, frunză, galben, gaură, iarbă, ierta, împinge, împlini, înapoi, început, înfunda, înger, lumina, luminare, lună, mai, mare, mut, nară, nici, noi, nor, paşti, pepene, petrece, piersic, ploaie, poamă, pulbere, putred, puţi, rîpă, ruga, sapă, sat, saţ, sămînţă, sănătos, seară, scaun, scălda, scîndură, scoarţă, seu, spinare, spînzura, stinge, stoarce, strămuta, tăciune, tei, tîrziu, trimite, tuna, uita 1, uita 2, vargă, vechi, vierme, vinde, visa, zău, zbate, zbor, zece . . . . . 3

440—564. acolo, acru, acum, ajutor, alerga, amiază, apoi, aprinde, aramă, aripă, ascuţi, atinge, atrage, baie, boteza, brumă, ca, capră, călare, căra, cărbune, căsca, căuş, chiar, cîmp, culege, culme, cuptor, deşert, doar, dor, ducă, fapt, farmec, feri, fir, foaie, freca, frig, frunte, fum, geană, ghem, ieri, iţă, îmbăta, împrumut, înălţa, încerca, închina, încotro, îndărăt, îneca, înfige, îngroşa, întîi, joi,



Scanned with OKEN Scanner

judeca, jug, lepădа, linge, lingură, linte, mas, măduvă, măr, mie, miel, miez, muc, nap, neted, nimeni, nou, nucă, orz, pară, păduche, pădure, păsat, piedică, pieri, piuă, plop, potîrniche, punte, pupa, putină, rădăcină, răpune, răpi, repede, ridiche, sac, sarcină, sănătate, săruta, sătura, scuipa, scutura, seară, sâmbătă, sorbi, speria, spic, spin, subsuoară, şi, teacă, teme, tindă, toartă, toca, tremura, tufă, uda, unt, urca, urî, usca, vecin, veghea, vis, zvînta . . . . . 2

565—725. acoperi, adăpost, adevăr, adică, adînc, aiurea, albie, argint, asemăna, asuda, asupra, aştepta, atare, aţîţa, bată, bătrîn, bea, beat, bucă, burete, buric, caş, cătinel, către, căţea, căţel, ceapă, cep, cerbice, codru, corb, curte, dator, de, depăna, descărca, descoperi, duminică, eu, făt, fierbe, fin, frîntură, fulger, fus, găleată, ghindă, greutate, iască, inel, împărţi, împunge, încăleca, încărca, încet, înceta, închega, îndesa, înfrunta, îngheţa, îngîna, înnoda, întîmpla, între, întuneca, învăţa, laţ, lăuda, lucra, luntre, lupta, lut, margine, martor, măciucă, mărăcine, măruntaie, meu, minţi, mişca, mînca, mînea, mîngîia, mînie, morişcă, muiere, mura, mustra, muta, negoţ, nimeni, oaste, opt, osînză, oua, pasăre, pată, păpuşă, părinte, păs, pătrat, pecingine, pescuitor, peţitor, pisa, plăcea, plumb, porumb, pradă, prună, pui, purcea, rază, rărunchi, răzbuna, revărsa, rîie, rîma, rînă, săgeată, sără, scînteie, scurge, seceră, secure, senin, sete, sine, singur, sîngera, smulge, soră, soţ, spată, spumă, spurca, strîmt, stup, subţia, sudoare, sulă, supăra, şa, şase, tămîie, tăun, ţînţar, undă, unghi, urgie, va, vacă, vineri, vioară, viperă, vită, vîrtos, vîrtute, vultur, zeu, zgîria . . . . . 1

Aşadar un număr de 725 de cuvinte din cele vreo 2 000, ce aparţin fondului latin primordial al limbii române, iau parte *activă* la formarea unor expresii şi locuţiuni. Am subliniat atributul pentru a evidenţia absenţa din lista de mai sus a multor elemente cu rol pasiv, în sensul că sînt fie simple cuvinte de legătură, fie unităţi lexicale folosite în accepţiunea lor uzuală, semnificaţia globală a expresiei depinzînd de



Scanned with OKEN Scanner

un alt component al acesteia. Este cazul, de pildă, al verbului *a face* în expresii ca *a-şi face mendrele, a face de petrecanie,* în care accentul semantic al sintagmei cade pe substantiv. Adesea termenul pe care l-am numit *pasiv* suportă substituirea cu altul apropiat, fără ca sensul global să fie sensibil afectat *(a netezi calea/drumul, a încălzi/creşte/ţine şarpele la sîn),* dar acest criteriu nu este neapărat decisiv în precizarea statutului cuvîntului, deoarece există situaţii cînd ambii constituenţi sînt înlocuibili *(a i se duce/a-i merge vestea/buhul),* ca să nu mai vorbim de altele în care tocmai termenul caracterizant, ce conferă plasticitate expresiei, poate suferi substituirea fără comutare semantică: *beat criţă/cuc/frînt/krup/lemn/lulea/mort/tei/tîlv/turtă/tun.*

Notînd cu *e* numărul de expresii la care un cuvînt dat participă ca „protagonist", în sensul observaţiilor de mai sus, repartiţia numerică a elementelor aparţinînd fondului latin moştenit în funcţie de valoare parametrului *e* este cea din tabelul 19.1.

*Tabelul 19.1.*

| $e \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 725 | 564 | 439 | 343 | 277 | 226 | 189 | 162 |
| ln (Nr. cuv.-bază) | 6,59 | 6,34 | 6,08 | 5,84 | 5,62 | 5,42 | 5,24 | 5,09 |

| $e \geq$ | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 149 | 132 | 116 | 107 | 98 | 86 | 76 | 66 |
| ln (Nr. cuv.-bază) | 5,00 | 4,88 | 4,75 | 4,67 | 4,58 | 4,45 | 4,33 | 4,19 |

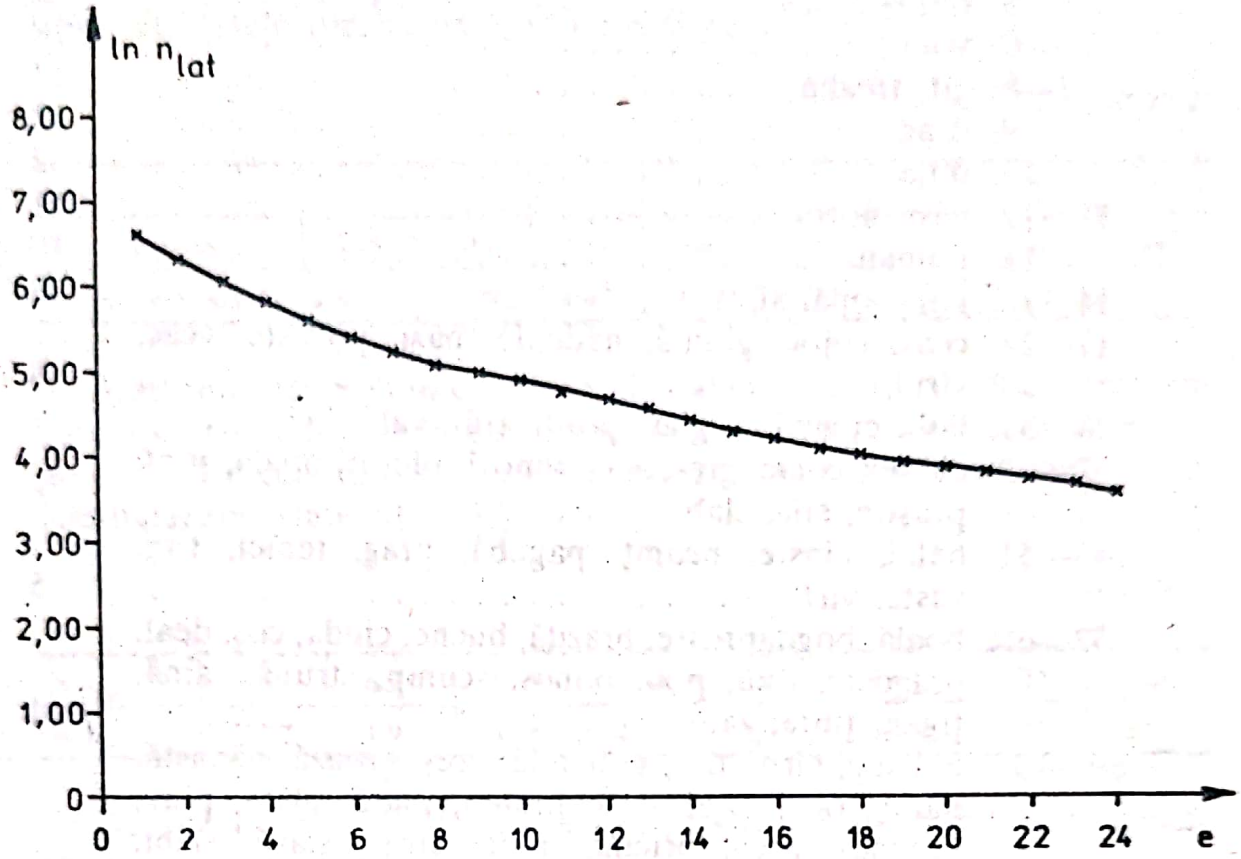
| $e \geq$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 59 | 57 | 52 | 48 | 45 | 42 | 41 | 35 |
| ln (Nr. cuv.-bază) | 4,08 | 4,04 | 3,95 | 3,87 | 3,81 | 3,74 | 3,71 | 3,56 |

Reprezentarea grafică a valorilor logaritmate din linia a treia a tabelului 19.1, în maniera, deja familiară cititorului, în care am făcut-o şi în cazul productivităţii şi polisemiei, evidenţiază neliniaritatea funcţiei respective în raport cu argumentul *e* (fig. 19.1). Prin urmare componenta latină se comportă, din punctul de vedere al participării la expresii, cam la fel ca sub aspectul numărului de sensuri ale cuvintelor şi într-un mod esenţialmente diferit de cel întîlnit mai înainte, în studiul productivităţii lexicale. Experienţa dobîndită în paragraful anterior ne-a învăţat să nu ne grăbim însă de a interpreta



Scanned with OKEN Scanner

FIG. 19.1




acest fapt ca pe o infirmare a „ipotezei liniarităţii". Ştim acum că neliniaritatea variaţiei distribuţiei unităţilor lexicale de o origine dată în raport cu un anumit parametru ales drept criteriu de clasificare nu constituie o piedică de principiu în calea existenţei unei relaţii liniare între această distribuţie şi cea referitoare la o altă componentă etimologică a vocabularului. Cazul polisemiei este lămuritor în acest sens.

Pentru a verifica în ce măsură lucrurile stau la fel şi în privinţa participării la expresii, vom examina pe rînd repartiţia sub acest aspect şi a altor aporturi lexicale importante. Începem cu elementul vechi slav. După numărul de expresii la care iau parte, cuvintele-bază de această origine se clasifică astfel:

| *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|
| 1. rînd | 29 |
| 2. drum | 26 |
| 3. nevoie | 25 |
| 4. vreme | 24 |

| | | |
|---|---|---|
| 5. | obraz | 23 |
| 6. | voie | 18 |
| 7—8. | gît, treabă | 14 |
| 9. | drag | 13 |
| 10. | milă | 12 |
| 11—12. | găsi, noroc | 11 |
| 13. | pomană | 10 |
| 14—16. | gol, popă, sfînt | 9 |
| 17—24. | ceas, cojoc, glumă, nădejde, păzi, poveste, veac, vîrî | 8 |
| 25—31. | bob, cumpănă, glas, poftă, trăi, val, vină | 7 |
| 32—42. | ciolan, colac, grai, leac, omorî, plesni, poală, praf, praştie, silă, slab | 6 |
| 43—51. | bălai, cinste, neamţ, pagubă, prag, temei, tîrg, veste, vîrf | 5 |
| 52—68. | boală, bogdaproste, brazdă, buche, ciudă, coş, deal, dragoste, duh, pod, ponos, scump, strună, taină, ţigan, ţintă, zare | 4 |
| 69—105. | belciug, cîrpi, dar, grămadă, greş, groază, năpastă, nărav, ocol, oţel, oţet, pinten, plasă, plată, plug, prăpădi, prea, pricină, rană, sfîrşi, slavă, slăbi, soroc, stîlp, străjă, suci, tabără, tablă, temelie, tîlc, topi, topor, trăsnet, trup, zăbavă, zid, zîmbi | 3 |
| 106—147. | beli, blid, ciocan, cîrpă, cleşte, cocoş, croi, drojdie, dungă, evanghelie, icoană, ispravă, istov, lipi, lopată, mănăstire, molcom, muncă, munci, oblu, obşte, opri, pară(= văpaie), pizmă, pomeni, porni, potrivă, povară, pravilă, price, rac, răscol, rod, roi, rufă, sită, slujbă trăsni, trebui, vîrstă, vodă, zori | 2 |
| 148.—226. | aleluia, amin, azbuche, bîrlog, boier, bold, buc, ceaslov, clei, clin, clipi, clopot, colind, corabie, dar (subst.), dîră, gleznă, glumi, greşi, gunoi, hîrtie, hram, hrăni, iordane, ispiti, iute, jalbă, jale, jăratec, lăstun, leat, leit, logodnă, meni, mironoşiţă, molie, nădrag, nemţesc, nicovală, obicei, orîndui, osie, ovăz, pacoste, pestriţ, pipăi, pir, pîrgă, plavie, plăti, pleavă, plocon, pofti, popas, post, prăpastie, razna, răzbi, rîcîi, sfară, slugă, sminti, smuci, stană, streaşină, tîrî, trezi, trudă, ulei, urni, veleat, vlădică, vrabie, vrednic, zadar, zălog, zdravăn, zmeu | 1 |



Scanned with OKEN Scanner

Tabelul 19.2. sistematizează, după modelul cu care cititorul s-a obişnuit, datele oferite de lista de mai sus.

*Tabelul 19.2.*

| $e \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 226 | 147 | 105 | 68 | 51 | 42 | 31 | 24 |
| ln (Nr. cuv.-bază) | 5,42 | 4,99 | 4,65 | 4,22 | 3,93 | 3,74 | 3,43 | 3,18 |

Alăturînd acum liniile ultime ale tabelelor 19.1 şi 19.2., obţinem coordonatele a 8 puncte din planul $ln\, n_{lat} - ln\, n_{sl}$, suficiente pentru a ne da o idee despre natura relaţiei dintre aceste două componente ale vocabularului (tabelul 19.3.).

*Tabelul 19.3.*

| $ln\, n_{lat}$ | 6,59 | 6,34 | 6,08 | 5,84 | 5,62 | 5,42 | 5,24 | 5,09 |
|---|---|---|---|---|---|---|---|---|
| $ln\, n_{sl}$ | 5,42 | 4,99 | 4,65 | 4,22 | 3,93 | 3,74 | 3,43 | 3,18 |

Reprezentate grafic în figura 19.2., punctele respective se grupează strîns în jurul dreptei de ecuaţie:

$$ln\, n_{sl} = 1{,}492\ ln\, n_{lat} - 4{,}421 \qquad (19.1.)$$

determinată, ca şi precedentele, cu ajutorul metodei celor mai mici pătrate.

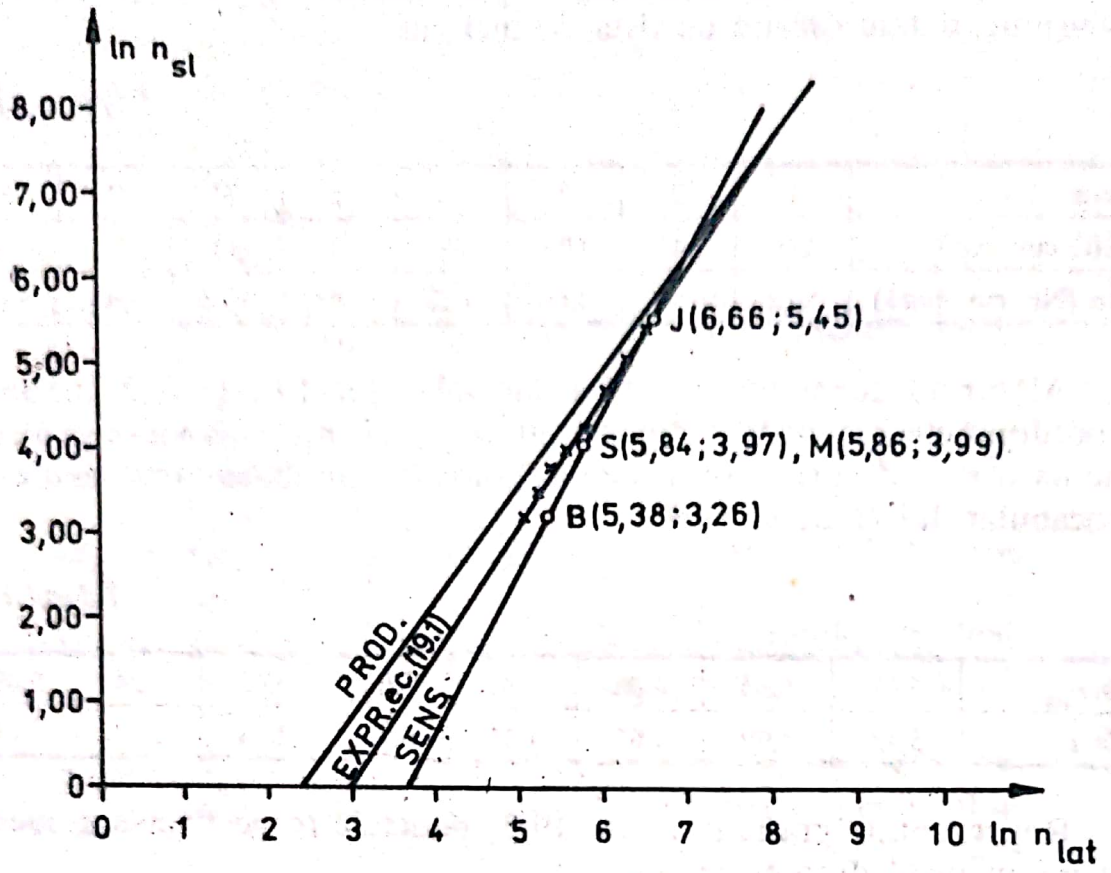
Ipoteza liniarităţii se confirmă aşadar şi în acest caz, impunîndu-se acum drept lege generală a raporturilor dintre fondul latin primordial al limbii române şi elementul vechi slav, atît din punctul de vedere al productivităţii lexicale cît şi din cel al numărului de sensuri şi din cel al bogăţiei de expresii asociate cuvintelor. Aceste trei relaţii liniare pot fi reprezentate sinoptic, figura rezultată vizualizînd raporturile latin moştenit-vechi slav sub toate cele trei aspecte considerate aici (fig. 19.2).

Se vede limpede că, privite din unghiul capacităţii de a intra în expresii, împrumuturile slave ocupă în raport cu cuvintele de origine latină o poziţie intermediară între aceea mai puternică pe care le-o conferă perspectiva productivităţii lexicale şi cea, sensibil inferioară, la care le reduce abordarea semantică. Faptul că, într-un domeniu numeric de o amplitudine echivalentă cu cea a unui dicţionar de cca. 50 000 de cuvinte, dreapta „productivităţii" (PROD) este situată



Scanned with OKEN Scanner

FIG. 19.2




pretutindeni deasupra celei corespunzătoare numărului de expresii (EXPR), iar aceasta, la rîndul ei, deasupra dreptei „polisemiei"(SENS) ne permite să conchidem că adoptarea productivităţii drept unic criteriu al importanţei cuvintelor flatează aportul slav, tot la fel cum o analiză exclusiv semantică îl dezavantajează în confruntarea cu fondul latin.

Fireşte, tripleta de criterii cu care am operat nu epuizează spectrul de parametri lingvistici ce definesc poziţia cuvintelor în sistemul vocabularului, dar nivelul ridicat de generalitate al modelului îngăduie extinderea perspectivei şi asupra altor, oricît de multe, caracteristici lexicale. Un exemplu la îndemînă ar fi cel al frecvenţei de folosire, indicator cam „urgisit" de noi, căruia îi datorăm, oricum, o reparaţie. Vom încerca să o oferim într-un capitol următor. Deocamdată menţionăm numai, anticipînd asupra rezultatelor ce vor fi prezentate mai tîrziu, posibilitatea de a înfăţişa pe o aceeaşi diagramă atît graficele productivităţii, polisemiei şi participării la expresii, cît şi datele furnizate de inventarele de frecvenţă pe care le-au întocmit diferiţi cercetători.

Pentru cazul raporturilor latin moştenit-vechi slav, dispunem de cîteva statistici, bazate pe corpusuri de structuri şi extinderi diferite.

Astfel, printre cele mai frecvente 522 de cuvinte din totalul de 4 547 unităţi numărate de colectivul condus de Valeriu Şuteu în eşantionul de 50 000 de cuvinte-text din beletristica secolului nostru la care ne-am mai referit (v. capitolul 2), un număr de 345 sînt de origine latină. Provenienţă slavă au, potrivit autorilor statisticii, 69 de cuvinte. Dintre acestea trebuie însă eliminate unităţile neoslave *(ba, bolnav, boltă, da, grădină, horă, iată, sărac)*, cîteva derivate formate pe teren românesc de la baze vechi slave *(căminar, privire, veşnicie)*, trei cuvinte a căror origine slavă nu este deloc sigură *(sută, vorbă, vorbi)* şi două care, în mod cert, nu sînt slave *(labă* < magh. *làb, lipsi* < ngr. *lipó)*. Rămîn 53 de termeni vechi slavi faţă de 345 latini, cuplu de valori care, după logaritmare, poate fi reprezentat pe diagrama din figura 19.2, sub forma unui punct de abscisă 5,84 şi ordonată 3,97. Notat cu *S*, acest punct cade aproape exact pe dreapta SENS ce descrie relaţiile latin-slav din punctul de vedere al polivalenţei semantice a cuvintelor (modelul $M_S$). El coincide practic cu un alt punct experimental *M* (5,86 ; 3,99) obţinut într-un mod asemănător din statistica lui C. Maneca, menţionată şi ea în capitolul 2. Între cele mai frecvente 1007 unităţi lexicale, din cele 6475 extrase din materialul pe care l-a excerptat (totalizînd 50 000 de cuvinte text), acest din urmă cercetător a găsit 349 de termeni latini moşteniţi şi 54 de cuvinte slave, adică aproape aceleaşi cifre ca şi V. Şuteu, deşi numărul total de cuvinte distincte a variat de la simplu la dublu (1007 faţă de 522)! Merită reţinută această invarianţă relativă a raportului latin-slav, în condiţiile în care componentele mai recente ale vocabularului înregistrează fluctuaţii enorme: 324 de termeni de provenienţă „franceză“ la C. Maneca, faţă de numai 3 (eventual 7, dacă includem şi neologismele etichetate drept „internaţionale“) la V. Şuteu.

Un alt punct, *B* (5,38; 3,26), ne este furnizat de inventarul de frecvenţe stabilit de Gh. Bolocan pe baza unor texte aparţinînd stilului publicistic (218 cuvinte latine şi 26 slave din cele mai frecvente 660). În ciuda valorilor absolute foarte diferite ale coordonatelor sale, acest punct se situează şi el foarte aproape de dreapta SENS, ca o confirmare a faptului că introducerea parametrului frecvenţă nu modifică sensibil imaginea pe care ne-am format-o pe baza celorlalţi indicatori asupra raportului de forţe dintre fondul latin şi componenta slavă veche a vocabularului limbii române.

Pentru a completa acest tablou, să mai considerăm o statistică, cea mai cuprinzătoare, dar şi cea mai controversată, cuprinsă în cunoscutul dicţionar de frecvenţe ale cuvintelor româneşti datorat lui Alphonse



Scanned with OKEN Scanner

Juilland. Analiza distribuţiei pe origini a unităţilor ce ocupă primele 1907 ranguri în lucrarea lexicografică menţionată a fost efectuată de Mihaela Bîrlădeanu în contextul studierii „vocabularului reprezentativ al limbilor romanice"[1]. Această cercetătoare a numărat 782 de cuvinte latine moştenite şi 233 de termeni vechi slavi. Exprimate logaritmic, aceste date constituie coordonatele unui punct *J* (6,66 ; 5,45), situat şi el în imediata apropiere (abaterea este de numai 2%) a dreptei SENS. Putem conchide că, cel puţin în ceea ce priveşte raporturile dintre stratul latin primordial şi împrumuturile vechi slave, uzajul se află într-o concordanţă atît de strînsă cu structura semantică încît statistica frecvenţelor de folosire a cuvintelor se prezintă drept aproape redundantă în raport cu ierarhia indusă de criteriul polisemiei.

Nu continuăm acest excurs, întrucît problematica uzajului merită o discuţie mai detaliată, pe care o amînăm pentru moment, spre a nu pierde din vedere obiectivul principal al capitolului de faţă. După examinarea componentelor latină şi slavă veche, să trecem în revistă situaţia participării la expresii a cuvintelor nederivate de origine „franceză":

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1—3. | idee, notă, onoare . . . . . | 6 |
| 4—9. | circulaţie, cont, parolă, raport, rezervă, termen . | 5 |
| 10—22. | act, atenţie, balanţă, bază, credit, egal, forţă, liber, natură, nerv, proces, respect, serviciu . . . . . | 4 |
| 23—52. | acord, baricadă, bord, cauză, caz, comandă, comun, discuţie, drapel, dura, evidenţă, exclusivitate, figură, grad, graţie, interes, mască, minut, moment, naţie, plus, poziţie, realitate, revoluţie, rigoare, scenă, simplu, situaţie, ton, valoare . . . . . | 3 |
| 53—120. | activ, acţiune, adresă, alfa, ansamblu, apel, ataca, auspiciu, centru, chit, clasă, comparaţie, consecinţă, consideraţie, contra, cor, corp, critică, discreţie, dispoziţie, dosar, efect, excepţie, formă, gardă, gaz, general, grup, iluzie, importanţă, impresie, indiferent, intenţie, intenţiona, justiţie, | |

[1] *Vocabularul reprezentativ al limbilor romanice* (coordonator: M. Sala), Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1988.



Scanned with OKEN Scanner

legalitate, libertate, lux, masă (⟨ fr. *masse*), memorie, modă, moral, morală, pană („de motor"), paradă, parcurs, pică, plan, placă, practică, prezenţă, publicitate, rampă, refuz, rezon, rond, sens, senzaţie, serie, siluetă, spectacol, stil, strict, teatru, teorie, termina, vigoare, viteză . . . . . 2

121—304. abstracţie, absurd, abundenţă, accent, achiziţie, acţiona, adeziune, afirmativ, aluzie, ambiţie, amfiteatru, amonte, anvergură, aparenţă, apariţie, artă, asentiment, aspect, atitudine, atmosferă, atracţie, audiţie, augur, aval, avangardă, avans, aviz, avocat, bagaj, balon, bancă, beneficiu, bloc, bronz, bucal, busolă, cantitate, chestiune, competenţă, concurs, condiţie, consistentă, constitui, conştiinţă, contact, contrabandă, contracta, contradictoriu, contribuţie, conţumacie, crocodil, curbă, curent, declin, declina, definiitiv, desuetudine, detaliu, diapazon, dimensiune, disponibil, distanţă, dramă, duzină, egidă, excelenţă, exerciţiu, extrem, extremă, favoare, fază, fizionomie, front, galerie, generaţie, gratis, homeric, instanţă, intern, intermediu, intimitate, jantă, literatură, localitate, manieră, marcă, mat, merit, miniatură, minoritate, mobiliza, monolit, monosilabic, motiv, muzeu, naftalină, navetă, necesitate, neglijeu, nivel, notorietate, nocturn, nul, numerar, obol, ocazie, oficiu, opĭnie, organic, original, origine, pachet, pantă, paralelă, parantează, pedală, periferie, permanenţă, persoană, perspectivă, pilulă, piramidă, pistol, planton, politică, post, prealabil, preferinţă, presiune, prezenta, prismă, priză, profit, progres, proporţie, public, puls, pupitru, radical, ramă, rang, rasă, relativ, replică, respiraţie, revelaţie, revers, rezultat, satisfacţie, secret, serios, seriozitate, sezon, societate, sos, stat (fr. *état*), statuie, subsidiar, substanţă, succesiune, surdină, şoc, şosea, tablou, tangent, telefon, temă, toaletă, total, trafic, tragic, trampă, travesti, tren, trenă, tribut, tur, unison, util, vervă, viager, vid, zel, zero . . . . . 1

Lista este destul de lungă, dar dacă se ţine seama de numărul uriaş de neologisme „franceze", ea ne apare drept mai degrabă modestă. Cele 304 cuvinte care intră în expresii nu reprezintă decît 2,01% din totalul de 15.121 termeni „francezi" nederivaţi din DLRM, pe cînd în cazul fondului latin, acelaşi raport atingea 40,23%. Pentru împrumuturile vechi slave valoarea este intermediară, 20,36%, rezultat de natură să sugereze existenţa unei relaţii direct proporţionale între vechimea afluenţilor vocabularului şi „producţia" lor de expresii şi locuţiuni. Dacă acceptăm şi de data aceasta (v. capitolul 12) că fondul lexical latin al limbii române datează de prin anul 150 şi că DEL, publicat în 1985, reflectă stadiul limbii române din jurul anului 1980, urmează că vîrsta medie ce trebuie luată în considerare într-o eventuală transpunere în termeni matematici a dependenţei de vechime a numărului de expresii şi locuţiuni este de circa 1830 de ani. Să presupunem că avem de a face cu o relaţie simplă, liniară, de forma:

$$p_i\,(\%) = a \cdot t_i \qquad (19.2)$$

unde $p_i\,(\%)$ este procentul reprezentat de cuvintele de origine $i$ care participă la expresii din totalul termenilor de aceeaşi provenienţă, iar $t_i$ — vechimea medie în limba română a cuvintelor de obîrşia respectivă. În cazul elementului latin, $p_{lat}\,(\%) = 40{,}23$ şi $t_{lat} = 1830$, valori care, introduse în ecuaţia (19.2), ne permit să calculăm coeficientul $a$:

$$a = 40{,}23 : 1830 = 0{,}02198 \text{ şi } 1/a = 45{,}49$$

Vîrsta medie a unui aport lexical caracterizat prin valoarea $p_i\,(\%)$ a procentului de cuvinte ajunse în expresii va fi atunci aproximată cu ajutorul formulei:

$$t_i = 45{,}49\,p_i \qquad (19.3)$$

Pentru elementul vechi slav, $p_{sl} = 20{,}36$ şi $t_{sl} = 926$ de ani, ceea ce revine la plasarea centrului de greutate al acestei influenţe pe la mijlocul veacului al XI-lea, iar pentru cel „francez", $p_{fr} = 2{,}01$ şi $t_{fr} = 91$ de ani, vîrstă corespunzătoare anului 1889. Ambele datări prezintă un grad ridicat de verosimilitate şi concordă remarcabil cu valorile determinate pe calea, cu totul diferită, a examinării producţiei de derivate: anii 975—1025 pentru aportul vechi slav şi 1880 pentru cel „francez".



Scanned with OKEN Scanner

Repartiţia neologismelor „franceze" în funcţie de numărul de expresii la care participă este înfăţişată în tabelul 19.4.

*Tabelul 19.4.*

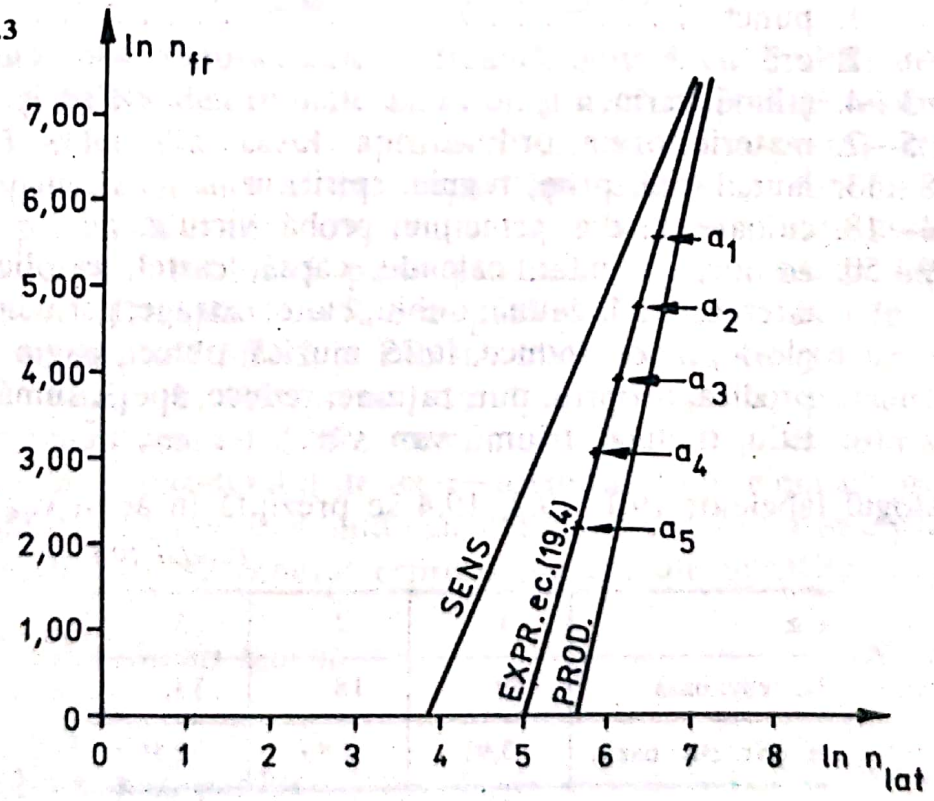
| $e \geq$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nr. cuv-bază | 304 | 120 | 52 | 22 | 9 |
| ln (Nr. cuv.-bază) | 5,72 | 4,79 | 3,95 | 3,09 | 2,20 |

Folosindu-ne şi de tabelul 19.1., avem la dispoziţie următoarele 5 puncte pentru deducerea alurii relaţiei latin-„francez": $a_1(6{,}59;\ 5{,}72)$, $a_2(6{,}34;\ 4{,}79)$, $a_3(6{,}09;\ 3{,}95)$, $a_4(5{,}84;\ 3{,}09)$, $a_5(5{,}62;\ 2{,}20)$. Reprezentate în figura 19.3, aceste puncte, aproape perfect coliniare, determină dreapta de ecuaţie:

$$ln\ n_{fr} = 3{,}588\ ln\ n_{lat} - 17{,}922 \qquad (19.4)$$

Ipoteza liniarităţii se vede aşadar confirmată şi în acest important caz. Observaţia, referitoare la elementul vechi slav, potrivit căreia participarea la expresii se supune unei legi de variaţie de acelaşi tip cu cea care guvernează producţivitate lexicală şi polisemia, dar ocupă

FIG. 19.3



un loc intermediar între aceste două extreme, își păstrează și aici valabilitatea, dar raporturile s-au inversat: după cum se observă în fig. 19.3., dreapta SENS se situează în cazul elementului „francez", deasupra dreptelor EXPR și PROD, în timp ce pentru aportul slav ea ocupa, dimpotrivă, poziția cea mai coborîtă. Investigarea expresiilor ne apare în consecință drept o cale de mijloc, foarte indicată pentru caracterizarea globală a raporturilor dintre constituenții etimologici ai vocabularului, tocmai întrucît legea de variație pe care ne-o descoperă mediază între normele de comportare a cuvintelor sub aspectul producerii de derivate și sub cel al expansiunii sensurilor.

Avînd în vedere asemănările existente între aportul lexical „francez" și cel latin savant, componente a cărora demarcare în DLRM e destul de convențională (evitarea diagnosticării dublei etimologii disjunge artificial aceste două clase, în fapt amplu intersectate, cum bine ne-o arată indicațiile, mult diferite, din DEX), ne așteptăm ca imaginea oferită de analiza influenței latine literare să prezinte analogii cu cea obținută pentru elementul neologic „francez". Clasificarea după numărul de expresii din DEL a cuvintelor-bază considerate în DLRM drept latine savante este următoarea:

| | *Unități lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | punct | 15 |
| 2. | oră | 6 |
| 3—4. | știință, termen | 5 |
| 5—7. | materie, ordin, ordine | 4 |
| 8—13. | laur, literă, prim, regulă, spirit, uz | 3 |
| 14—18. | culoare, medic, principiu, probă, virtute | 2 |
| 19—50. | ad-hoc, calendar, calende, capsă, castel, catolic, circumstanță, daună, dubiu, etate, extrage, germen, glorie, index, induce, linie, muzică, obiect, pagină, predica, propriu, pur, rațiune, reduce, speță, sumă, titlu, traduce, triumf, van, văl | 1 |

Analogul tabelelor 19.1, 19.2, 19.4 se prezintă în acest caz astfel:

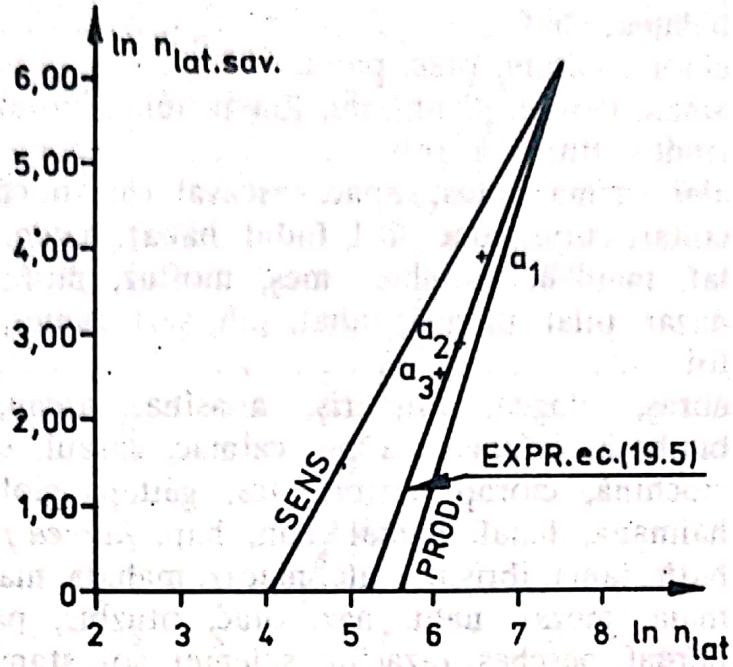
*Tabelul 19.5*

| $e \geq$ | 1 | 2 | 3 |
|---|---|---|---|
| Nr. cuv.-bază | 50 | 18 | 13 |
| ln (Nr. cuv.-bază) | 3,91 | 2,89 | 2,56 |



Scanned with OKEN Scanner

FIG. 19.4




Din pricina numărului relativ redus de elemente, această clasă ne furnizează numai 3 puncte, $a_1$(6,59; 3,91), $a_2$(6,34; 2,89) şi $a_3$(6,09; 2,56), nesituate întrutotul coliniar, dar printre care metoda celor mai mici pătrate ne permite să trasăm dreapta de ecuaţie:

$$ln\ n_{lat.sav.} = 2{,}732\ ln\ n_{lat} - 14{,}201 \qquad (19.5)$$

Figura 19.4 alătură această dreaptă, notată cu EXPR, dreptelor PROD şi SENS determinate anterior şi a căror poziţie relativă este similară celei din cazul aportului „francez". Aspectul semantic precumpăneşte şi aici asupra capacităţii lexicogene, fenomen firesc dacă îl privim în lumina observaţiei că în ambele situaţii avem de a face cu termeni neologici, puţin productivi, însă bogaţi în înţelesuri de „import". Dreapta EXPR e dispusă şi de această dată în unghiul format de dreptele SENS şi PROD, mai aproape de cea de a doua, fapt ce poate fi interpretat drept o dovadă a asemănării de comportament general dintre elementul „francez" şi cel latin savant.

Un număr considerabil de expresii conţin drept element pivot cîte un cuvînt de origine turcă. Iată lista celor 110 unităţi lexicale provenite din turceşte care au generat expresii consemnate în DEL:

| *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|
| 1. zor . . . . . | 11 |
| 2—3. haz, perdea . . . . . | 7 |

| | | |
|---|---|---|
| 4. | para . . . . . | 6 |
| 5—6. | balama, chef . . . . . | 5 |
| 7—10. | chior[1], haram, olac, papuc . . . . . | 4 |
| 11—21. | aman, ciorbă, divan, hal, liman, lulea, miraz, oca, rindea, tinichea, turc . . . . . | 3 |
| 22—52. | alai, basma, belea, capac, caşcaval, catran, cherem, cîntar, cutie, edec, fitil, fudul, habar, haide, iureş, laf, maidan, merchez, meş, mofluz, moft, nart, nazar, pilaf, pingea, rahat, şah, şart, tabiet, tătar, toi . . . . . | 2 |
| 53—111. | abraş, alagea, alişveriş, anasîna, arcan, bei, bimbaşă, cafenea, caftan, caimac, caraul, chimir, ciochină, ciorap, ciripie, fes, găitan, giol, hac, haimana, halal, hamal, han, hap, *harcea-parcea,* hatîr, iama, ibrişin, leafă, macaz, mahala, mahmur, maia, mezat, naht, naz, otac, otuzbir, palavră, perdaf, peşcheş, razachie, selemet, soi, stamboală, surghiun, şiret, tain, tarabă, tavă, taxim, temenea, tichie, tighel, toptan, topuz, veresie, zapt, zeflemea | 1 |

*Tabelul 19.6*

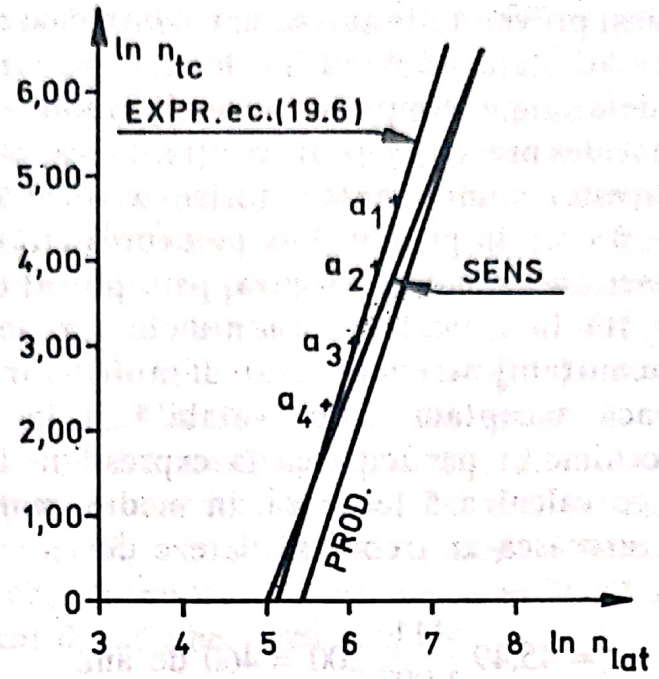
| $e \geq$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nr. cuv.-bază | 111 | 52 | 21 | 10 |
| ln (Nr. cuv.-bază) | 4,71 | 3,95 | 3.04 | 3.30 |

Sintetizat în tabelul 19.6, bilanţul elementului turcesc poate fi transpus grafic (fig. 19.5) sub forma a 4 puncte, $a_1$(6,59; 4,70), $a_2$(6,34; 3,95), $a_3$(6,09; 3,04), $a_4$(5,84; 2,30), aliniate remarcabil pe dreapta:

$$ln\, n_{tc} = 3{,}205\ ln\ n_{lat} - 16{,}421 \qquad (19.6)$$

[1] De fapt, cel puţin într-una dintre expresiile reţinute de DEL, nu a fost vorba la origine de împrumutul *chior* (⟨ tc. *kör*), ci de adjectivul moştenit *chiar* (⟨ lat. *clarus*), confundat cu cel dintîi în momentul cînd, ieşit din circulaţie, acesta din urmă nu mai supravieţuia decît în unităţi frazeologice ca *apă chi(o)ară.* E totuşi greu de spus dacă şi în *para chioară* avem de a face tot cu sensul de „pur şi simplu" al lat. *clarus* sau la o aluzie la un eventual orificiu practicat în monedă, expresia însemnînd deci „ban găurit", după cum nu putem şti cu certitudine nici dacă *noroc chior* provine dintr-un mai vechi *noroc chiar* sau constituie doar o referire glumeaţă la cecitatea Fortunei.

FIG. 19.5




Spre deosebire de cazurile studiate anterior, fasciculul alcătuit din dreptele PROD, SENS şi EXPR este de data aceasta mult mai „adunat". Explicaţia fenomenului trebuie căutată în vîrsta medie nici prea mare nici foarte scăzută a împrumuturilor turceşti. Producţia totală de derivate fiind o funcţie de timp, dreapta PROD se situează la stînga dreptei SENS în cazul influenţelor mai vechi şi la dreapta ei în cazul celor mai recente. O vîrstă mijlocie se traduce printr-o mai mare apropiere, ce poate merge pînă la coincidenţă, a celor două drepte. Inclusă, în general, în domeniul delimitat de acestea, dreapta EXPR va forma împreună cu ele un mănunchi strîns, împrejurare fericită pentru cercetător, căruia i se oferă astfel şansa de a putea caracteriza aproape univoc locul ocupat în limbă de componenta etimologică respectivă: din orice punct de vedere am privi lucrurile (din perspectiva productivităţii, a polisemiei, ori a numărului de expresii), raporturile cu fondul latin moştenit vor fi cam aceleaşi, ceea ce nu se poate, în schimb, afirma nici despre constituenţii mai vechi (substratul, elementul vechi slav) şi nici despre cei de dată recentă (neologismele romanice).

Oarecum ieşită din comun în raport cu comportarea scontată pe baza observaţiilor de pînă acum este abundenţa de expresii care conţin cuvinte de origine turcă. Urmărind traseul dreptei EXPR, se constată că ea părăseşte de timpuriu domeniul cuprins în unghiul format de dreptele PROD şi SENS şi urcă deasupra acestora, cotele

înalte pe care le atinge evidenţiind o particularitate interesantă a materialului lexical provenit din turceşte: cuvintele de această origine supravieţuiesc nu atît datorită familiilor lexicale pe care au izbutit să şi le constituie de-a lungul timpului sau mulţimii de sensuri pe care le exprimă, cît mai ales prezenţei lor în locuţiuni şi expresii. Rezistenţa la eroziune a acestei componente istorice a vocabularului limbii române se bazează deci, în primul rînd, pe expresivitatea cuvintelor turceşti, pe capacitatea lor de a asigura, prin plusul de plasticitate pe care o prezintă în raport cu sinonimele lor de alte origini, acel caracter memorabil strict necesar dăinuirii lor în conştiinţa vorbitorilor. Dacă acceptăm drept valabilă şi în cazul de faţă relaţia dintre vechime şi participarea la expresii în forma dată de expresia (19.3), se calculează lesne că, în medie, materialul lexical de provenienţă turcească ar trebui să dateze de:

$$t = 45{,}49\,\frac{111}{1\,097}\,100 = 460 \text{ de ani.}$$

adică din prima jumătate a veacului al XVI-lea, rezultat ce diferă apreciabil de determinările noastre anterioare, care plasaseră centrul de greutate al influenţei turceşti tocmai la mijlocul secolului al XVIII-lea. O concluzie fermă este greu de tras, avînd în vedere faptul că formula utilizată mai sus constituie numai o simplă conjectură, întemeiată pe analiza unui fond de date fatalmente incomplet (inventarul de expresii din DEL e departe de a fi exhaustiv). Dacă însă valabilitatea relaţiei 19.3 s-ar confirma, am putea conchide că numărul de împrumuturi turceşti intrate în expresii este cam de două ori mai mare decît cel calculat în ipoteza localizării grosului importurilor de termeni din limba turcă în intervalul 1750—1760. Într-adevăr, aceasta ar corespunde unei vîrste medii de cca. 220—230 de ani, faţă de anul de referinţă 1980, valoare care introdusă în expresia 19.3, furnizează $p = 4{,}84 \div 5{,}06\%$, adică exact jumătate din procentul real dedus din DEL: $p = \frac{111}{1\,097} = 10{,}12\%$. Concluzia ar fi că numărul termenilor împrumutaţi din turceşte care au dat naştere la expresii este considerabil superior celui ce ar corespunde vîrstei medii, relativ scăzute, a acestor elemente.

Să nu ne pripim însă acordînd acestei teze mai mult credit decît merită, căci lucrurile pot fi privite şi altfel. Funcţia $t_i = 45{,}49\,p_i$ depinde aparent de un singur parametru, dar în realitate conţine două



Scanned with OKEN Scanner

argumente, deoarece indicatorul $p_i$ reprezintă raportul a două valori autonome: numărul cuvintelor de origine *i* care au generat expresii şi numărul total de cuvinte de aceeaşi provenienţă. Or, acest din urmă parametru nu a rămas constant de-a lungul timpului. A existat, fără îndoială, o vreme, nu prea îndepărtată, cînd vocabularul limbii române cuprindea mult mai multe cuvinte de origine turcă decît în zilele noastre. Cele 111 unităţi lexicale pătrunse în expresii erau, la vremea constituirii expresiilor respective, exponentele unei mulţimi de împrumuturi turceşti ce depăşea cu mult cele 1097 de cuvinte care mai dăinuiau în limbă în momentul redactării DLRM. Presupunînd, de exemplu, că în jurul anului 1850 marea majoritate (să zicem 80%) a expresiilor conţinînd cuvinte provenite din turceşte erau deja constituite, ceea ce este destul de plauzibil, şi menţinînd data de 1750 pentru vechimea medie a componentei turceşti a vocabularului, putem calcula cu ajutorul formulei 19.3 că, pe la mijlocul secolului trecut, materialul lexical de origine turcă totaliza:

$$n_{tc} = \frac{111 \times 0{,}80 \times 45{,}49 \times 100}{100} = 4\,040 \text{ de cuvinte}$$

adică de aproape 4 ori mai multe decît astăzi. Avînd în vedere ipotezele destul de arbitrare ce au stat la baza calculului nostru, rezultatul obţinut contează mai puţin prin valoarea sa absolută, pe care o alegere diferită a valorii parametrilor consideraţi o poate face să crească sau să scadă în limite foarte largi, cît prin fenomenul pe care ni-l semnalează: o presiune lexicală efemeră, ca cea exercitată de vremelnica influenţă turcească, poate lăsa urme mai durabile în domeniul formaţiilor ce depăşesc limitele cuvîntului (locuţiuni, expresii, zicale, proverbe) decît în vocabularul propriu-zis. Odată trecut valul înrîuririi respective, termenii ce nu s-au dovedit în stare să-şi creeze o familie lexicală sau o armătură semantică puternică sînt îndepărtaţi din limbă, ceea ce se întîmplă mult mai anevoie cu unităţile participante la expresii, chiar dacă, în destule cazuri, şi expresiile se perimează. În cîmpul turcismelor, nu puţine sînt situaţiile în care cuvinte învechite, care altminteri ar fi dispărut demult, supravieţuiesc exclusiv graţie unor expresii: a dormi cît un *bei,* a trăi ca *bimbaşa,* a face *harcea-parcea,* a rămîne (lăsa) *mofluz,* a fi la *cheremul,* a scoate la *mezat,* cu *anasîna,* cu *otuzbirul,* cu *toptanul* etc. Dincolo de speculaţiile matematice, care, în absenţa unor puncte de sprijin suficient de ferme, s-ar putea dovedi un exerciţiu falacios, rămîne în



Scanned with OKEN Scanner

afara oricărei îndoieli faptul că, spre deosebire de aportul vechi slav, reliefat cel mai pregnant prin ierarhizarea cuvintelor după productivitate, aportul neologic („francez" şi latin savant) apare potenţat cu precădere de o analiză semantică, iar contribuţia turcească — de cercetarea patrimoniului de expresii şi locuţiuni al limbii române.

Din acest punct de vedere, elementul grecesc ar putea prezenta asemănări cu materialul lexical de obîrşie turcească. Cuvintele provenite din greceşte ce figurează în colecţia de expresii din DEL sînt:

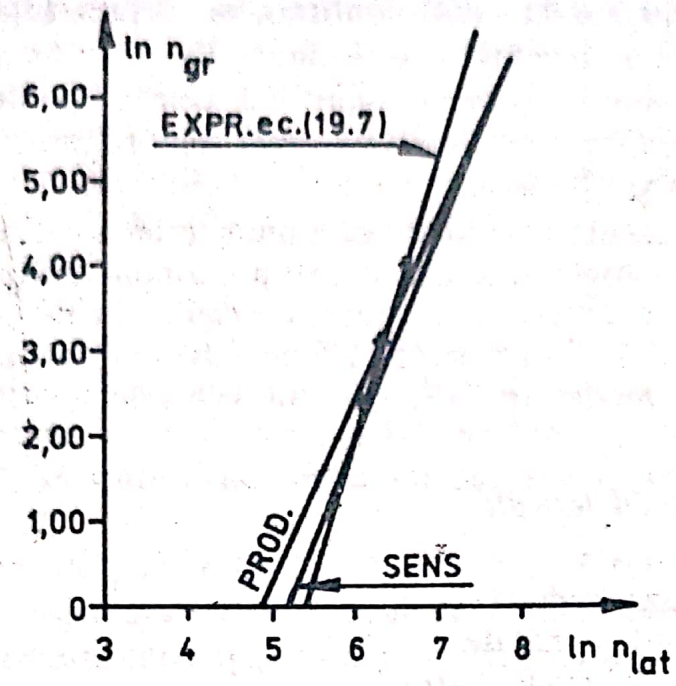
| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1—2. | buzunar, piper . . . . . | 5 |
| 3—6. | condei, mutră, sfoară, tipar . . . . . | 4 |
| 7—11. | calimeră, mirosi, noimă, pastramă, zodie . . . . . | 3 |
| 12—22. | aghios, boboc, ieftin, lipsi, oftică, paraxin, prochimen, sigur, splină, stradă, ţîfnă . . . . . | 2 |
| 23—55. | ananghie, calapod, catastif, cort, dîrmon, doxă, fasole, folos, honipsi, ifos, iotă, ison, nostimadă, oleacă, omidă, partidă, patos, piftie, pirostrie, prisos, proaspăt, rapanghel, saramură, sardea, săpun, sosi, spanac, stambă, stemă, stimă, stomac, tiflă, zahăr . . . . . | 1 |

Acceptînd şi de data aceasta condiţia autoimpusă de a nu lua în considerare, în prelucrarea informaţiei statistice, decît mulţimile ce cumulează cel puţin 10 cuvinte, lista de mai sus, se materializează în tripleta de puncte $a_1$(6,09; 2,40), $a_2$(6,34; 3,09), $a_3$(6,59; 4,00). Ecuaţia dreptei EXPR duse printre aceste puncte, în aşa fel încît suma pătratelor distanţelor dintre ele şi dreaptă să fie minimă, este:

$$ln\, n_{gr} = 3{,}293\ ln\, n_{lat} - 17{,}715 \qquad (19.7)$$

În reprezentarea sinoptică din figura 19.6, dreapta EXPR ni se înfăţişează cu o pantă mai abruptă decît dreptele PROD şi SENS, situaţie întîlnită şi în cazul elementului turcesc. De data aceasta însă, două dintre cele trei puncte folosite la trasarea dreptei, $a_1$ şi $a_2$, sînt amplasate în interiorul unghiului format de dreptele determinate anterior, ceea ce nu se întîmpla în figura 19.5. Faptul trebuie

FIG. 19.6




interpretat drept o consecinţă a vechimii mai mari a infuenţei greceşti, răspunzătoare de deplasarea către stînga a dreptei PROD, căci, aşa cum am mai spus-o, cu cît un aport lexical este mai vechi, cu atîta el a generat mai multe derivate. Deosebindu-se de componenta turcă a vocabularului prin deplasarea accentului dinspre polisemie către productivitate, elementul grecesc se comportă, în compensaţie, într-un mod oarecum similar în ceea ce priveşte participarea cuvintelor la expresii şi locuţiuni, în sensul că panta dreptei EXPR este şi aici mai mare decît cele ale dreptelor PROD şi SENS. În cazul elementului turcesc, această caracteristică se asocia cu un excedent absolut al termenilor efectiv întîlniţi în raport cu numărul estimat pe baza aplicării ecuaţiei 19.3 la intervalul de timp furnizat de metoda de datare cu ajutorul productivităţii. Cum se prezintă din acest punct de vedere situaţia împrumuturilor greceşti?

În capitolul 12 s-a calculat cu relaţia 12.6 că centrul de greutate al îndelungatei influenţe greceşti cade probabil în perioada 1570—1590. Luînd ca termen de referinţă anul 1980, vechimea medie a unităţilor provenite din limba greacă ar fi aşadar de cca. 400 de ani. Formula 19.3 ne permite atunci să evaluăm la aproximativ 400 : 45,49 = 8,79% numărul împrumuturilor greceşti care ar fi fost de aşteptat să fi pătruns în expresii. Faţă de cele 624 de unităţi de această proveniență existente în DLRM, este deci probabil ca vreo 55 să participe la expresii. Prezenţa a exact 55 de cuvinte în lista termenilor greceşti identificaţi



Scanned with OKEN Scanner

în expresiile cuprinse în DEL confirmă pe deplin supoziţia de mai sus, consolidîndu-ne încrederea în formula 19.3 care ne-a furnizat pînă acum datări aflate în perfect acord cu cele obţinute din studiul productivităţii pentru trei dintre cele mai importante influenţe: slavă veche, grecească şi „franceză".

Dată fiind prezenţa neîndoielnică a unor termeni de substrat printre cuvintele cu etimologie necunoscută, vom examina mulţimea unităţilor lexicale comune cu albaneza în directă corelare cu colecţia elementelor de origine obscură. Următoarele 27 de cuvinte-bază participante la expresiile consemnate de DEL au un echivalent atestat în limba albaneză:

| *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|
| 1—2. buză, căciulă | 7 |
| 3—4. gard, jumătate | 6 |
| 5—7. moş, traistă, vatră | 5 |
| 8—11. ceafă, groapă, grumaz, rînză | 4 |
| 12—18. gata, ghimpe, mal, mireasă, pălărie, pupăză, raţă | 3 |
| 19—23. curma, cursă, mînz, şopîrlă, ţap | 2 |
| 24—27. brîu, grapă, mătrăgună, murg | 1 |

în timp ce alte 99 sînt de origine necunoscută:

| | |
|---|---|
| 1. vorbă | 45 |
| 2. băga | 24 |
| 3. ban | 10 |
| 4. pînză | 9 |
| 5. bot | 8 |
| 6—7. băţ, gheară | 7 |
| 8. scăpăra | 6 |
| 9—12. brînză, burduf, sută, toană | 5 |
| 13—15. butuc, reteza, stăpîn | 4 |
| 16—27. baltă, bîrzoi, burtă, clanţă, gaie, hăţ, mămăligă, neagă, pungă, rîcă, rîşniţă, sperlă | 3 |
| 28—50. bîtă, bîzdîc, blat, boacă, boţ, bumben, buştean, colb, dodie, fuior, fulg, gărgăun, goangă, melc, mîrcă, moţ, neghină, poznă, prăjină, preş, toană, ţanc, ţîţînă, zbrancă | 2 |
| 51—99. bleau, boacănă, bocnă, buc, buh, buiestru, catrafuse, chelfăneală, chitic, dar (conj.), dihai, | |

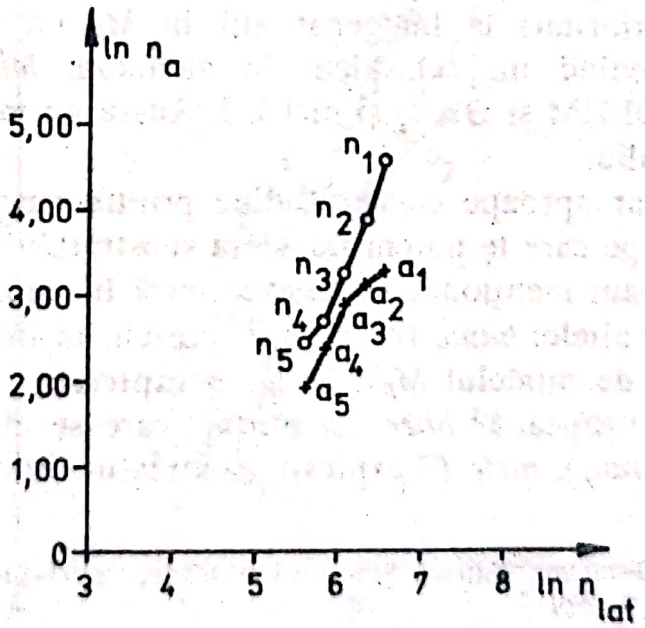
fiţă, ghes, gogoaşă, hachiţe, hăbuc, hău, landră, lela, leu, mau, mendre, nasuliu, năzar, oarzăn, perie, pîlnie, praftură, prohab, răbda, refec, resteu, rezema, rostogol, străin, şorici, şparli, talaghir, teşilă, teşmeni, ticlui, tîlv, ţîră, ţîţ, ţoşcă, ţuşcă, zbîrli, zburda . . . . . 1

Am reprezentat laolaltă, pe o aceeaşi diagramă (fig. 19.7), punctele $a_1$(6,59; 3,30), $a_2$(6,34; 3,14), $a_3$(6,09; 2,90), $a_4$(5,84; 2,40), $a_5$(5,62; 1,95), corespunzătoare materialului lexical comun cu albaneza, şi punctele $n_1$(6,59; 4,60), $n_2$(6,34; 3,91), $n_3$(6,09; 3,30), $n_4$(5,84; 2,71), $n_5$(5,62; 2,48), aferente listei cuvintelor de proveniență obscură, prilej de a observa că:

1. Mulţimea de puncte $\{a_i \mid i = 1, 2, 3, 4, 5\}$ formează o linie frîntă (A), ceea ce sugerează că inventarul de cuvinte care a stat la baza determinării ei este incomplet. Apare drept foarte probabilă prezenţa unui număr oarecare de termeni de substrat ascunşi printre unităţile cu etimologie necunoscută.

2. Mulţimea de puncte $n_i$ este dispusă şi ea sub forma unui unghi, dar de orientare diametral opusă în raport cu cel alcătuit de linia frîntă (A). Această complementaritate geometrică constituie un argument în plus în favoarea tezei că ceea ce lipseşte mulţimii

FIG. 19.7



cuvintelor comune cu albaneza este de căutat în mulţimea elementelor de origine obscură.

Într-adevăr, punctul $n_5$ e situat deasupra prelungirii dreptei $\overline{n_1n_2n_3n_4}$, în timp ce punctul $a_5$ este amplasat cu mult sub prelungirea dreptei $\overline{a_1a_2a_3}$, dovadă că există în lista cuvintelor de proveniență necunoscută un excedent de cuvinte prezente în cel puţin 5 expresii care corespunde unui deficit de termeni cu aceeaşi proprietate din lista elementelor autohtone. Confruntarea listei dintîi cu repertoriul cuvintelor „sigure sau probabil autohtone (preromane) care lipsesc în albaneză" întocmit de I. I. Russu[1], precum şi cu mulţimile de cuvinte pe care modelele matematice precedente (cel al productivităţii $M_p$ şi cel al polisemiei $M_s$) le recomandau ca posibile elemente de substrat ne semnalează că:

a. Dintre termenii participanţi la 5 sau mai multe expresii, verbului *băga* şi substantivelor *brînză, pînză* şi *gheară* li s-au avansat etimologii preromane.

b. Trei dintre cuvintele de mai sus rezultaseră drept susceptibile de a aparţine substratului şi din discutarea ecuaţiilor productivităţii şi polisemiei. Astfel, printre termenii ce posedă mari şanse de a se dovedi autohtoni, modelul $M_p$ ni-i indică pe *pînză* şi *brînză* (v. paragraful 10.17), iar modelul $M_s$ pe acelaşi *pînză* şi pe *băga* (capitolul 16).

c. Un al cincilea cuvînt căruia toate metodele de pînă acum converg în a-i atribui origine geto-dacă este *burduf,* termen figurînd în listele candidaţilor prioritari la indigenat atît în $M_p$ cît şi în $M_s$ şi, pe deasupra, posedînd un echivalent în albaneză, *burdhë*[2], deşi nici dicţionarele (DLRM şi DEX) şi nici I. I. Russu nu menţionează acest important detaliu.

Există aşadar aproape cu certitudine printre cuvintele de origine obscură cîteva pe care le putem transfera substratului fără mari riscuri de a greşi. Le-am menţionat pe cele ce intră în minimum 5 expresii, dar mai sînt şi altele: *butuc* (4 expresii; prezent în lista lui I. I. Russu şi recomandat de modelul $M_p$), *burtă* (3 expresii; a fost apropiat de rădăcina indoeuropeană[3] *bher-* „a purta," care se află probabil şi la originea lui *burduf*), *melc* (2 expresii; candidatură sprijinită şi de $M_s$),

[1] *Limba traco-dacilor,* Editura Ştiinţifică, 1967, pp. 215—216.
[2] I. L. R. II, p. 329.
[3] ILR II, p. 355.



Scanned with OKEN Scanner

*ghes, răbda, rezema, şoric, zburda* (cîte o expresie; toate socotite de I. I. Russu drept autohtone). Cum se modifică linia (A) în ipoteza adăugării acestor cuvinte la colecţia elementelor provenite din substrat?

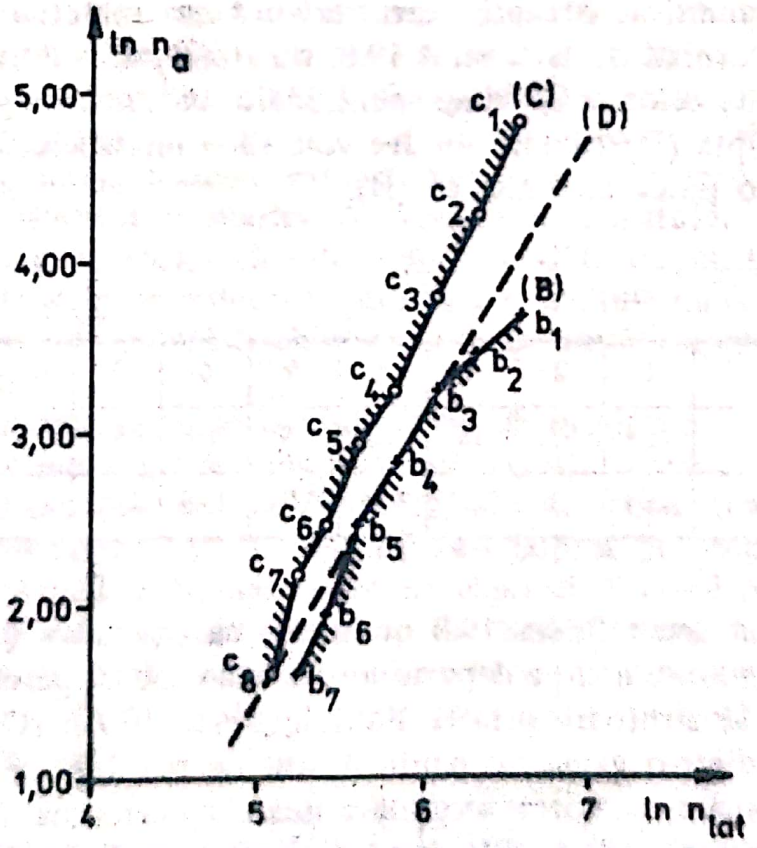
Bilanţul unităţilor autohtone are în acest caz aspectul din tabelul 19.7.

*Tabelul 19.7.*

| c ≥ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 41 | 32 | 26 | 17 | 12 | 7 | 5 |
| ln (Nr. cuv.-bază) | 3,71 | 3,47 | 3;26 | 2,83 | 2,48 | 1,95 | 1,61 |

Reprezentate în fig. 19.8., de data aceasta la o scară ceva mai mare, cele 7 puncte furnizate de tabelul 19.7 alcătuiesc tot o linie frîntă (B), semn că din colecţia de cuvinte acum îmbogăţită mai lipsesc încă mulţi termeni proveniţi din substrat. Într-un paragraf precedent (10.17) am încercat, întemeindu-ne pe concluziile desprinse din analiza

FIG. 19.8



productivităţii lexicale, să determinăm o limită superioară a numărului de cuvinte autohtone păstrate în românește. A rezultat cu acest prilej că volumul relictelor respective ar putea depăşi cu greu 185 de cuvinte. Pornind acum de la datele pe care ni le procură studiul expresiilor, vom încerca să oferim şi o evaluare a limitei inferioare a aceluiaşi parametru.

În lumina celor discutate mai sus, ne aşteptăm ca, asemenea tuturor celorlalte afluxuri lexicale istoriceşte circumscrise, influenţa substratului să se supună şi ea unei legi liniare. Printr-un transfer convenabil de unităţi din clasa cuvintelor cu etimologie necunoscută, linia frîntă (B) poate fi transformată într-o dreaptă. Această operaţie se poate efectua însă în nenumărate feluri. Există totuşi două limitări de care trebuie ţinut seamă: pe de o parte, dreapta căutată nu poate coborî în nici un punct al ei sub linia (B), iar pe de alta ea nu se poate ridica nici oricît de sus, dat fiind că rezerva de cuvinte disponibile pentru a fi adăugate la zestrea substratului nu este nemărginită. Cumularea celor două liste anterioare, cea a elementelor probabil autohtone şi cea a termenilor de origine obscură, ne permite să determinăm frontiera superioară a domeniului în care poate să se situeze prezumtiva dreaptă caracteristică a materialului lexical preroman. Pornind de la tabelul 19.8, care prezintă valorile rezultate din însumarea celor două clase menţionate, se construieşte linia (C) pe care dreapta (D) căutată nu are voie să o intersecteze nicăieri, la fel cum nu o putea face nici cu (B).

*Tabelul 19.8*

| $e \geq$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr. cuv.-bază | 126 | 73 | 45 | 26 | 19 | 12 | 9 | 5 | 4 | 3 |
| ln (Nr. cuv.-bază) | 4,84 | 4,29 | 3,81 | 3,26 | 2,94 | 2,48 | 2,20 | 1,61 | 1,39 | 1,10 |

Examinînd acum figura 19.8 observăm că libertatea de mişcare a dreptei (D) nu este nici pe departe atît de mare pe cît părea la început. Obligată să se strecoare printre limitele rigide (B) şi (C), ea are de ales numai dintr-o gamă de pante destul de restrînsă. Poziţia ce ne interesează aici e cea corespunzătoare înclinării celei mai mici posibile a dreptei, deoarece acesta este cazul care comportă cel mai mic număr



Scanned with OKEN Scanner

total de termeni autohtoni. Se vede pe figură că dintre toate variantele compatibile cu constrîngerile menţionate, cea care prezintă panta minimă este dreapta ce trece prin punctele $b_3$ şi $b_5$. Ecuaţia ei:

$$ln\ n_{substrat} = 1{,}660\ ln\ n_{lat} - 6{,}847 \qquad (19.8)$$

ne furnizează pentru abscisa $ln\ n_{lat} = 6{,}59$, corespunzătoare numărului total de cuvinte de origine latină prezente în expresii şi locuţiuni, ordonata $ln\ n_{substrat} = 4{,}09 = ln\ 60$. Trebuie să existe, prin urmare, în limba română cel puţin 60 de vocabule autohtone intrate în expresii. Pe de altă parte însă, am avut ocazia să constatăm mai înainte că fracţiunea pe care o reprezintă cuvintele ce ajung în expresii din totalul unui aport lexical dat depinde de vechimea influenţei respective şi că acest parametru nu poate în nici un caz depăşi valoarea de 40%, cît se înregistrează pentru fondul primordial latin. E de presupus că şi substratul tinde către o cifră apropiată. Oricum, chiar dacă rămîne sub această valoare, pragul de 40% constituie o limită cu ajutorul căreia putem estima numărul minim de cuvinte preromane care au supravieţuit pînă în zilele noastre. El este $n_{substrat} = 60{:}40\% = 150$. Coroborînd acest rezultat cu cel obţinut din studiul productivităţii, putem afirma acum, cu destulă siguranţă, că zestrea totală de termeni de substrat a limbii române cuprinde între 150 şi 185 de cuvinte.

Probabilitatea depăşirii, în ambele sensuri, a acestui interval este foarte mică, rezultat important dacă îl privim în perspectiva constatării că „potrivit părerilor exprimate de diverşi specialişti elementele de substrat din limba română variază între o sută şi aproape o mie"[1]. Reducerea drastică a marjei de variaţie a aportului autohton ne conduce către concluzia că cele vreo 160—170 de cuvinte pe care I. I. Russu le atribuie influenţei geto-dace constituie nu doar un nucleu reprezentativ, ci, probabil, chiar totalitatea unităţilor preromane din limba română.

Încheiem analiza materialului lexical de origine necunoscută cu o observaţie privind proveniența controversatului cuvînt *vorbă*. Graţie celor 45 de expresii la care participă, acesta ocupă o poziţie aparte în lista termenilor obscuri, devansîndu-l copios pe secundantul său, verbul *băga*, prezent în numai 24 de expresii. Studiul productivităţii lexicale ca şi cel al polisemiei lăsau deschisă alternativa apartenenţei lui *vorbă* la substrat sau la fondul vechi slav (v. paragraful 10.17 şi capitolul 16). Ce ne poate spune în această privinţă cunoaşterea numărului de expresii?

---

[1] A. Vraciu *Limba daco-geţilor* Ed. Facla, 1980, p. 165

Ştim deja că ecuaţiile modelului $M_s$ permit o evaluare a rangului probabil al celui mai bogat în expresii cuvînt de provenienţă dată. Astfel, admiţînd drept valabilă pentru substrat ecuaţia 19.8 se calculează uşor, prin anularea membrului drept al ecuaţiei, valoarea $ln\, n_{lat}$ la care e de aşteptat să survină primul cuvînt autohton în clasificarea cuvintelor româneşti după numărul de expresii şi locuţiuni pe care le-au generat. Rezultă $ln\, n_{lat} = 4{,}125$, adică $n_{lat} = 62$. Supunînd aceleiaşi proceduri ecuaţia 19.1:

$$ln\, n_{slav} = 1{,}492\, ln\, n_{lat} - 4{,}421$$

găsim poziţia probabilă a celui mai activ participant la expresii dintre cuvintele vechi slave: $ln\, n_{lat} = 2{,}963$ şi $n_{lat} = 19$.

Concluzia spre care ne conduce acest calcul comparativ este că, dacă printre primele 60—70 de cuvinte bogate în expresii există două de origine încă neidentificată, atunci sînt toate şansele ca termenul mai bine plasat în ierarhie să fie de origine slavă veche, iar celălalt, clasat pe un loc mai modest, să provină din substrat. Or, consultînd lista de la începutul acestui paragraf, am constatat deja că în zona respectivă se află exact două unităţi de origine obscură: *vorbă* (rangul 15, nr. expresii = 45) şi *băga* (rangul 39—41, nr. expresii = 24). Poziţia lor relativă constituie un argument puternic în favoarea originii salve a celui dintîi, ipoteză sprijinită şi de elementele invocate de diferiţi lingvişti în favoarea originii autohtone a lui *băga* (printre ei, cum s-a văzut, se numără acum şi autorul acestor rînduri, în măsura în care concluziile modelului $M_s$ converg asupra acestui punct cu opinia lui I. I. Russu[1]).

Dar, fiindcă sîntem pe terenul ipotezelor, să ne fie permis a mai avansa una. Avînd în vedere poziţia fruntaşă a substantivului *vorbă* în toate cele trei clasificări ale cuvintelor considerate pînă aici, singura componentă a vocabularului care ar mai putea disputa elementului vechi slav acest controversat cuvînt este... fondul latin moştenit!

Într-adevăr, valorile numerice ale celor trei indicatori de importanţă a cuvintelor îl situează pe *vorbă* în plin domeniu latin: dacă extragem din listele prezentate anterior termenii aflaţi simultan în toate cele trei clasamente pe unul din primele 100 de locuri, constatăm că ele sînt în totalitate moştenite, cu o unică (eventuală) excepţie care ar fi tocmai *vorbă*. Iată înşiruirea acestor cuvinte, avînd înscrisă în paranteze tripleta alcătuită din numărul de derivate, numărul de sensuri şi

[1] *Op. cit.* pp. 215—216.

numărul de expresii ale fiecăruia: *bate* (22, 20, 46), *bun* (34, 28, 30), *cruce* (15, 11, 40), *face* (20, 45, 46), *faţă* (34, 13, 23), *ochi* (25, 17, 82), *parte* (23, 9, 29), *pune* (20, 15, 52), *sta* (17, 16, 38) *strînge* (16, 10, 16), *trage* (21, 30, 39), *ţine* (17, 46, 84), *veni* (19, 12, 30), *vorbă* (17, 11, 45). Tot nu departe de limita primelor 100 de cuvinte se situează *vorbă* şi derivatul său *vorbi* în principalele inventare de frecvenţă. Primul ocupă rangurile 119—121 la V. Şuteu şi 139 la A. Juilland, iar cel de al doilea — rangurile 122—126 la V. Şuteu şi 100 la A. Juilland. Aceste poziţii pot părea modeste, dar, dacă, se ţine seamă de predominarea absolută a cuvintelor ajutătoare asupra celor cu sens lexical deplin în partea superioară a clasamentului, cuvîntul *vorbă* ne apare situat pe un loc privilegiat: dintre toate substantivele de origine nelatină, numai două *(copil* şi *vreme)* îl preced în lista de frecvenţe a lui A. Juilland. Statistica lui V. Şuteu le mai adaugă acestora pe *glas* şi *drum,* ceea ce nu schimbă esenţial situaţia. Oricum *vorbă* rămîne unul dintre cuvintele fundamentale ale limbii române, comparabil ca importanţă lingvistică numai cu cei mai însemnaţi dintre termenii latini moşteniţi. Ba mai mult decît atît: în elita celor 14 protagonişti lexicali enumeraţi mai sus, numai arbitrariul înşiruirii alfabetice îl plasează pe *vorbă* pe ultimul loc. De fapt însă, parametrii săi lingvistici îl îndreptăţesc să aspire la o poziţie mai bună, pe care am putea-o evalua astfel: cele trei numere ataşate fiecărui cuvînt pot fi privite drept coordonatele unui punct din sistemul cartezian tridimensional ale cărui axe sînt ataşate, respectiv, productivităţii, polisemiei şi numărului de expresii; în acest caz, raza vectoare a punctului, a cărei expresie matematică e dată de formula:

$$r = \sqrt{d^2 + s^2 + e^2}, \text{ unde}$$

$d$ = numărul de derivate
$s$ = numărul de înţelesuri
$e$ = numărul de expresii şi locuţiuni

constituie o măsură a importanţei cuvîntului. Ordonate după valorile descrescătoare ale lui $r$, cele 14 cuvinte menţionate prezintă următoarea ierarhie:

| CUVINTE | $r$ | CUVINTE | $r$ | CUVINTE | $r$ |
|---|---|---|---|---|---|
| 1. ţine | 97,3 | 6. trage | 53,5 | 11. parte | 38,1 |
| 2. ochi | 87,4 | 7. bun | 53,3 | 12. veni | 37,5 |
| 3. face | 67,4 | 8. *vorbă* | 49,3 | 13. strînge | 24,7 |
| 4. pune | 57,7 | 9. sta | 44,6 | 14. cruce | 23,9 |
| 5. bate | 54,8 | 10. faţă | 43,1 | | |



Scanned with OKEN Scanner

Se vede că termenul ce ne interesează suportă remarcabil comparaţia cu cele mai „multilaterale" cuvinte româneşti[1], clasîndu-se al doilea dintre substantive şi făcînd figură de excepţie absolută printre împrumuturi, avînd în vedere că apare încadrat exclusiv de cuvinte de origine latină.

E posibil să descindă şi el dintr-un etimon latin? Pentru a încerca să răspundem la această întrebare, vom semnala mai întîi faptul că, potrivit evoluţiei fonetice normale, pluralul *uerba* al lat. *uerbum* ar fi trebuit să conducă la un rom. *vărbă* (cf. lat. *uerus* > rom. *văr* care atestă dezvoltarea la incipit a semiconsoanei *u* în fricativa labiodentală sonoră *v*, precum şi transformarea *e* > *ă* în contextul *u-r;* pentru conservarea grupului *rb* cf. evoluţia lat. *barba* > rom. *barbă*), care printr-un fenomen deloc neobişnuit de disimilare vocalică ar fi putut da pe *vorbă.* Singurul punct neclar al acestei etimologii este considerarea unui plural ca bază de generare pentru singularul românesc. Ar trebui să presupunem o folosire independentă, destul de intensă şi bine delimitată semantic a pluralului *uerba,* care să fi determinat tratarea sa de către vorbitori drept un substantiv de tipul *pluralia tantum* în momentul cînd înţelesul, diferit, al singularului a fost preluat de lat. *conuentus* > rom. *cuvînt.* În această privinţă, ştim că *uerba* cunoştea şi un sens specific şi anume cel de „vorbe goale" (cf. la Horaţiu *uirtutem uerba putas* = „socoteşti virtutea o baliverna" şi la Cicero *uerba sunt atque ineptiae* = „sînt palavre şi nerozii", de unde şi expresia *uerba dare* cu sensul de „a amăgi", „a păcăli", „a înşela"). Reminiscenţele acestei deprecieri semantice (din care nu se împărtăşeşte şi sinonimul *cuvînt,* tratat, de regulă fără pic de ironie: *om de cuvînt, în toată puterea cuvîntului)* sînt detectabile în expresii româneşti ca *vorbă să fie!, ce atîta vorbă?,. nici vorbă!, mai încape vorbă?, nu-i vorbă!,* toate conţinînd mai mult sau mai puţin aluzia, de posibilă sorginte latină, la vacuitatea *vorbelor* în contrast cu *cuvintele.* În ipoteza, plauzibilă după opinia noastră, că specializarea semantică a pluralului l-a salvat pe *uerba* în timp ce concurenţa lui *conuentus* îl

---

[1] În sensul unei plurivalenţe, manifestate prin capacitatea de a genera deopotrivă cuvinte, înţelesuri şi expresii noi, triplă calitate pe care am evidenţiat-o reţinînd numai termenii ce posedă concomitent minimum 14 derivate, 9 sensuri şi 14 expresii. Să nu uităm să subliniem însă că pot există cuvinte neincluse în lista noastră care, din pricina hipertrofierii unuia dintre aceşti trei parametri, să înregistreze valori ale lui *r* superioare unora dintre cele reţinute anterior, cum ar fi de exemplu, *da* (13, 36, 71) cu *r* = 80,7, *cap* (12, 12, 69) cu *r* = 71 ,1, *lua* (5, 23, 65) cu *r* = 69,1 şi alte cîteva, aparţinînd fără excepţie fondului latin primordial al vocabularului.



Scanned with OKEN Scanner

înlătura din limbă pe *uerbum,* controversatul *vorbă* ar veni să se adauge fondului latin moştenit, epurînd de ultima prezenţă alogenă nucleul lexical cel mai intim al limbii romîne.

Această digresiune odată încheiată, reluăm firul analizei, continuînd cu examinarea situaţiei cuvintelor cu dublă etimologie bulgaro-sîrbocroată. Am mai avut ocazia de a menţiona dificultăţile separării nete a împrumuturilor paleoslave de cele neoslave. O parte considerabilă dintre cei 180 de termeni cărora DEX le atribuie două etimoane, unul bulgăresc şi altul sîrbocroat, pot aparţine stratului vechi slav, ipoteză sprijinită de datarea cu ajutorul formulei 12.6, bazate pe corelaţia dintre vechime şi productivitatea lexicală. Într-adevăr, pantele dreptelor caracteristice pentru elementul vechi slav ($m = 0{,}312$) şi pentru împrumuturile „bulgaro-sîrbocroate" ($m = 0{,}329$) fiind de valori apropiate, vîrsta celor din urmă rezultă numai cu un secol inferioară vîrstei celor dintîi: $t_{bg.scr.} = 832 - 878$ de ani faţă de $t_{sl} = 925 - 975$ de ani. Raportîndu-ne la momentul 1950, e vorba aşadar cam de anul 1100 în primul caz şi de anul 1000 în cel de al doilea.

Dintre cele 180 de cuvinte-bază zise „bulgaro-sîrbocroate", următoarele 34 au generat expresii consemnate în DEL:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | necaz | 11 |
| 2. | răboj | 7 |
| 3—4. | ba, colţ | 5 |
| 5. | năvală | 4 |
| 6—13. | ciupercă, clacă, piron, pîrlea, ploscă, potcoavă, tărîţă, tigaie | 3 |
| 14—20. | lele, păstra, pîrpăli, slănină, sticlete, sucală, tuli | 2 |
| 21—34. | baştină, bobot, busuioc, chiseliţă, claie, gaiţă, grădină, oişte, pilă, prepeliţă, surlă, tigvă, ţelină, zobi | 1 |

Procentual, ele reprezintă 18,89%, adică o proporţie nu cu mult inferioară celei înregistrate de împrumuturile slave vechi (20,36%). Pentru valoarea respectivă, relaţia 19.3 estimează la $t_{bg.scr.} = 45{,}49 \cdot 18{,}89 = 859$ de ani vechimea medie a termenilor cu dublă etimologie faţă de $t_{slav} = 926$, o diferenţă chiar puţin mai mică decît cea furnizată de metoda de datare bazată pe corelaţia productivitate-vechime.



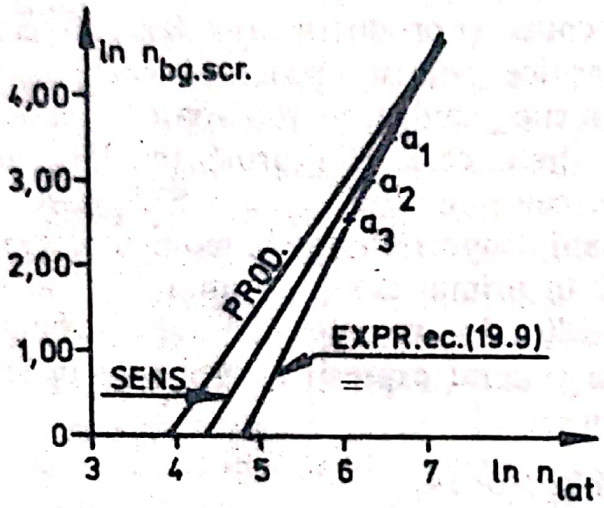
Scanned with OKEN Scanner

Punctele $a_1$(6,59; 3,53), $a_2$(6,34; 3,00), $a_3$(6,09; 2,56) care reflectă raporturile cantitative dintre componenta bulgaro-sîrbocroată şi cea latină sînt aproape perfect coliniare (fig. 19.9), determinînd dreapta:

$$ln\ n_{bg.scr.} = 1{,}976\ ln\ n_{lat} - 9{,}498 \qquad (19.9)$$

situată, precum se vede, dedesubtul dreptelor PROD şi SENS. Caracterul compozit al materialului lexical considerat ne împiedică să tragem vreo concluzie din această constatare.

FIG. 19.9

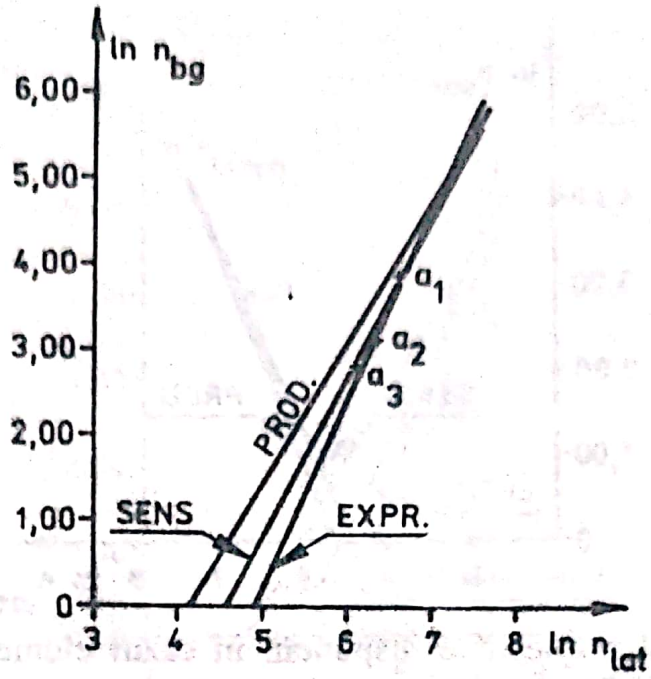



Lista cuvintelor-bază cu etimon unic bulgăresc intrate în expresii se prezintă astfel:

| | Unităţi lexicale | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | opincă | 9 |
| 2—3. | grijă, papară | 7 |
| 4–6. | rudă, slobod, şagă | 5 |
| 7—8. | pită, pustiu | 4 |
| 9—17. | cîrlig, gîrlă, gîscă, horă, matcă, miji, nisip, obială, sărac | 3 |
| 18—23. | covrig, dîrvală, hîrzob, păstra, praz, tîrcol | 2 |
| 24—49. | breaz, ciubăr, cloci, copcă, creangă, cremene, curcă, daiboj, dîrdoră, grozav, haină, japcă, jir, lanţ, lesne, leucă, meliţă, nadă, nălucă, odihni, sloi, sorcovă, suveică, verif, vopsea, zăpor | 1 |

Din totalul de 309 termeni de aceeaşi provenienţă număraţi de noi în DLRM, aceştia reprezintă 15,86%, procent corespunzător, potrivit

FIG.19.10




formulei 19.3, unei vechimi $t_{bg} = 45{,}49 \cdot 15{,}86 = 721$ de ani, adică aproximativ anului 1260 (faţă de data de referinţă 1980). Estimarea pe temeiul productivităţii indicase o perioadă ceva mai recentă: anii 1305—1340. Ecuaţia corespunzătoare aportului bulgăresc are forma:

$$ln\, n_{bg} = 1{,}780\, ln\, n_{lat} - 7{,}743 \quad \text{(v. fig. 19.10)}$$

Materialul lexical de origine exclusiv sîrbocroată prezent în expresii este, tot apud DEL, următorul:

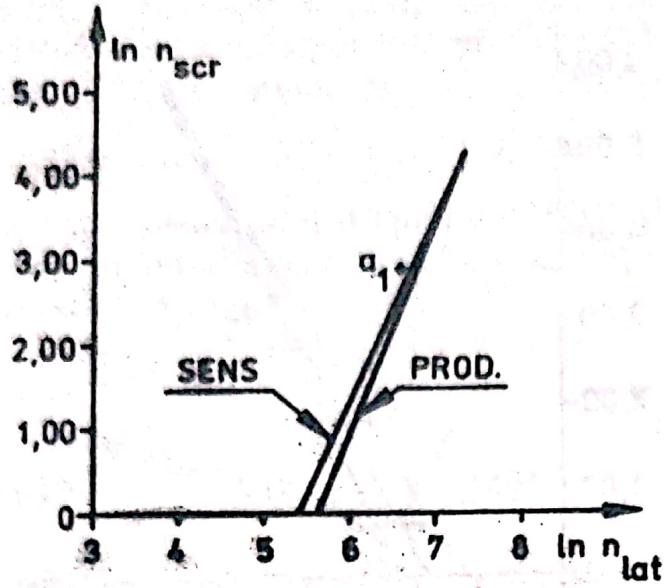
| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | pahar . . . . . | 4 |
| 2. | cîrd . . . . . | 3 |
| 3—5. | cîrcă, părăsi, pîrtie . . . . . | 2 |
| 6—18. | aliman, boltă, coţcă, duşcă, îndruga, ogar, platcă, pleaşcă, scai, stative, tipic, tîrnosi, zgomot . . . . | 1 |

Insuficientă pentru a defini un grafic, fie el liniar sau nu, această colecţie de cuvinte determină punctul $a_1(6{,}59;\ 2{,}89)$, situat în vecinătatea unghiului format de dreptele PROD şi SENS. (fig. 19.11). Nu sîntem deci în măsură să construim o ecuaţie a raporturilor numerice dintre elementul latin şi cel sîrbocroat, dar buna concordanţă dintre cele două drepte cunoscute, aproape coincidente, precum şi amplasarea punctului $a_1$ în imediata lor proximitate, subliniază comportarea unitară a elementului sîrbocroat din toate punctele de vedere luate în discuţie.



Scanned with OKEN Scanner

FIG. 19.11




De ceva mai multe date dispunem în cazul elementului maghiar. Pe baza listei cuvintelor intrate în expresii:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | gînd | 22 |
| 2. | chip | 14 |
| 3. | fel | 11 |
| 4. | tobă | 10 |
| 5. | neam | 8 |
| 6—7. | şatră, vileag | 4 |
| 8—11. | labă, marfă, modru, pildă | 3 |
| 12—19. | bai, bardă, bănat, gazdă, guler, ham, teleleu, ţarcă | 2 |
| 20—47. | aldămaş, bir, burzului, cătană, conci, dric, făgaş, fedeleş, hang, helge, hîrdău, hotar, lacăt, măgan, mertic, meşter, mistui, muştar, ocnă, oraş, panglică, poghiaz, pecetlui, pont, răvaş, sîrg, tontoroi, vamă | 1 |

se construieşte dreapta EXPR (fig. 19.12) de ecuaţie:

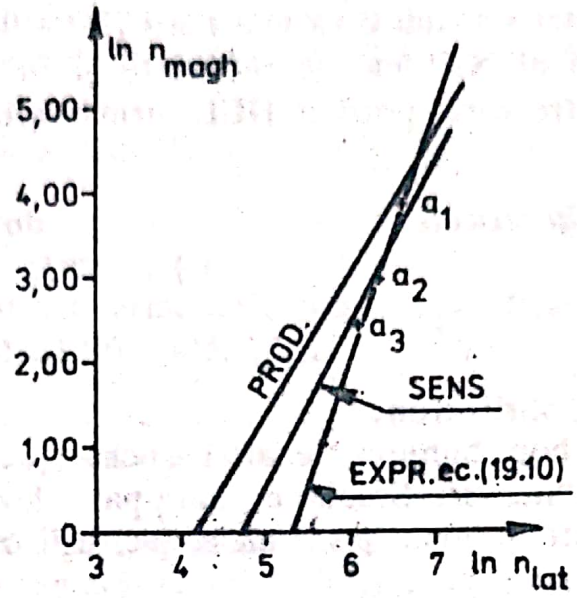
$$ln\ n_{magh} = 3{,}000\ ln\ n_{lat} - 15{,}947 \qquad (19.10)$$

Trebuie remarcat faptul că dispunerea reciprocă a celor trei drepte, PROD, SENS şi EXPR, este, şi de această dată, analoagă cu cea din cazurile aporturilor grecesc, bulgăresc şi bulgaro-sîrbocroat, ceea ce ar putea constitui un indiciu asupra vîrstei oarecum comparabile a acestor aporturi. Într-adevăr, în toate situaţiile menţionate e vorba de influenţe



Scanned with OKEN Scanner

FIG. 19.12




suficient de vechi pentru ca unităţile în cauză să-şi fi creat familii puternice de derivate, fapt ce se traduce prin amplasarea dreptei PROD la stînga celorlalte două, dar, în acelaşi timp, şi îndeajuns de recente pentru ca dreapta EXPR să nu fi putut depăşi încă, în lenta sa mişcare către stînga, aliniamentul dreptei SENS (fenomen întîlnit, în schimb, în cazul componentei vechi slave a vocabularului). Raportate la totalul de 476 de cuvinte de origine maghiară din DLRM, cele 47 de unităţi participante la expresii reprezintă o cotă procentuală de 9,87%. Datarea cu ajutorul formulei 19.3 furnizează un rezultat, la prima vedere, derutant: $t_{magh} = 449$ de ani, ceea ce corespunde plasării centrului de greutate al influenţei maghiare în jurul anul 1530. Responsabilă de această împingere exagerată către actualitate a unui aport pe care avem toate motivele de a-l considera mai vechi este prezenţa în dicţionar a unui număr mare de elemente regionale de vîrstă, în general, inferioară celei a împrumuturilor răspîndite în întreg spaţiul carpato-danubian. Cu totul altfel se prezintă lucrurile dacă ne referim numai la acestea din urmă: rămîn doar 300 de cuvinte de origine maghiară, dintre care 41 au intrat în expresii (din lista de mai sus doar 6 termeni şi anume *bai, bănat, conci, helge, măgan* şi *modru* sînt regionalisme). Procentul este de 13,67% şi el ne conduce la $t_{magh} = 622$ de ani, adică la o poziţionare a baricentrului influenţei ungureşti pe la jumătatea secolului al XIV-lea, rezultat nu cu mult diferit de cel obţinut prin metoda datării pe baza producţiei de derivate.

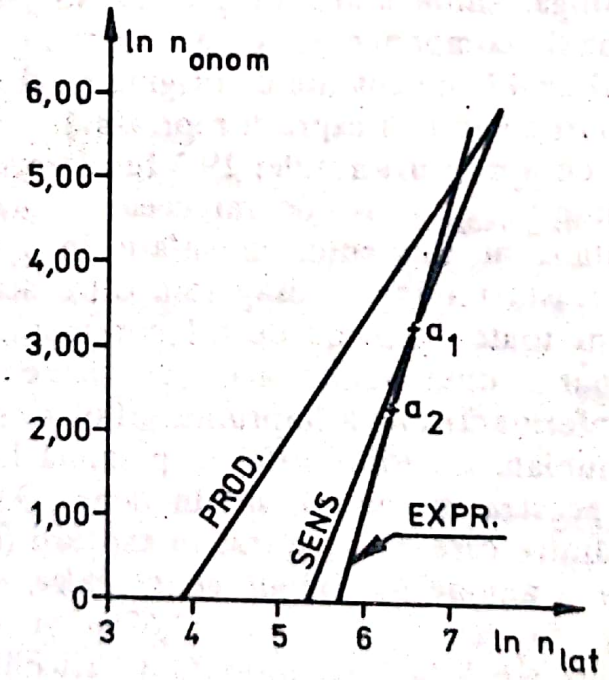
Aceeaşi amplasare relativă a dreptelor PROD, SENS şi EXPR (fig. 19.13), cu o deplasare mai accentuată către stînga a dreptei PROD, semn al unei vechimi ceva mai mari (formula 12.6 le atribuie în medie secolului al XIII-lea) se întîlneşte şi în cazul formaţiilor onomatopeice, dintre care, potrivit DEL, următoarele au pătruns în expresii:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | mîţă . . . . . | 7 |
| 2. | pic . . . . . | 5 |
| 3—5. | dura, sfîrîi, tronc . . . . . | 3 |
| 6—10. | cîrc, hop, huhurez, leoarbă, şfichi . . . . . | 2 |
| 11—26. | bîz, chiu, cîr, dîrdală, ei, hara-para, hîr, hîrţi, hîţ, hodoronc, huţa, pis, tralala, ţac, ţîţîi, zbughi . . | 1 |

Ecuaţia dreptei EXPR, trasată prin punctele $a_1$(6,59; 3,26) şi $a_2$(6,34; 2,30) este:

$$ln\ n_{onom} = 3{,}750\ ln\ n_{lat} - 21{,}464 \qquad (19.11)$$

**FIG. 19.13**




De fapt, unul dintre punctele utilizate se situează chiar pe dreapta SENS, iar abaterea celui de al doilea e atît de mică încît nici nu putem fi siguri că dreptele EXPR şi SENS nu coincid. Materialul lexical de

care dispunem este prea sărac pentru a ne permite să tragem concluzii categorice, observaţie în şi mai mare măsură valabilă pentru acele aporturi, încă mai deficitare, care nu permit nici măcar definirea unei drepte, fie ea oricît de aproximativă. Este cazul elementului german:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | şurub . . . . . | 5 |
| 2—4. | ort, şreang, ţel . . . . . | 2 |
| 5—14. | abstract, armonică, chix, clapă, faliment, pantof, stofă, şnur, şură, ţaglă . . . . . | 1 |

al celui italian:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1—4. | adio, asalt, monedă, piaţă . . . . . | 2 |
| 5—16. | anticameră, basta, brio, genere, medalie, papagal, paşaport, reuşi, revistă, speze, tapet, vulcan . . . | 1 |

sau al celui rusesc:

| | *Unităţi lexicale* | *Nr. expresii (DEL)* |
|---|---|---|
| 1. | poştă . . . . . | 4 |
| 2. | talie . . . . . | 3 |
| 3—4. | găluşcă, pensie . . . . . | 2 |
| 5—11. | bragă, holeră, leică, niznai, poliţie, smirnă, stroi | 1 |

FIG. 19.14

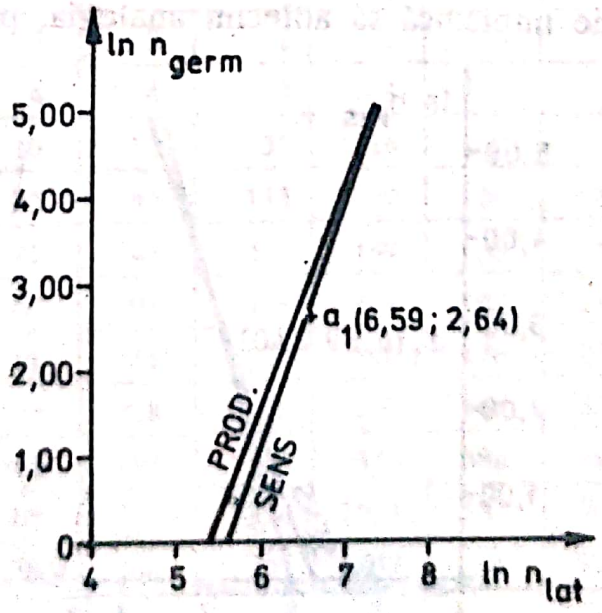
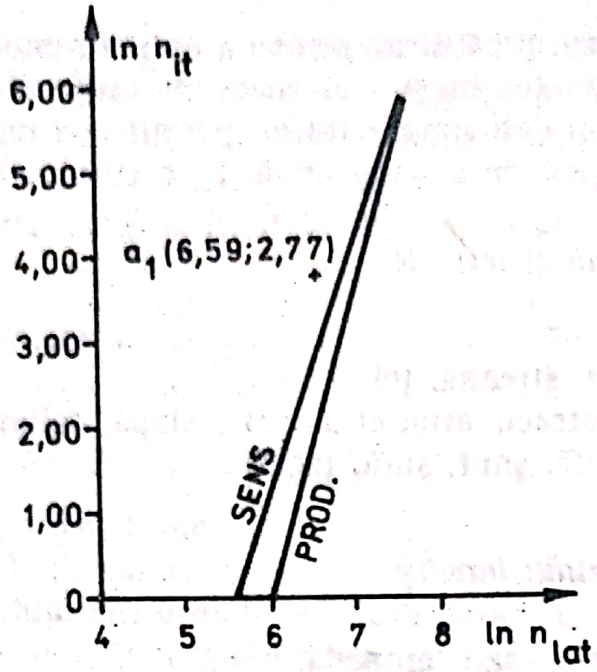


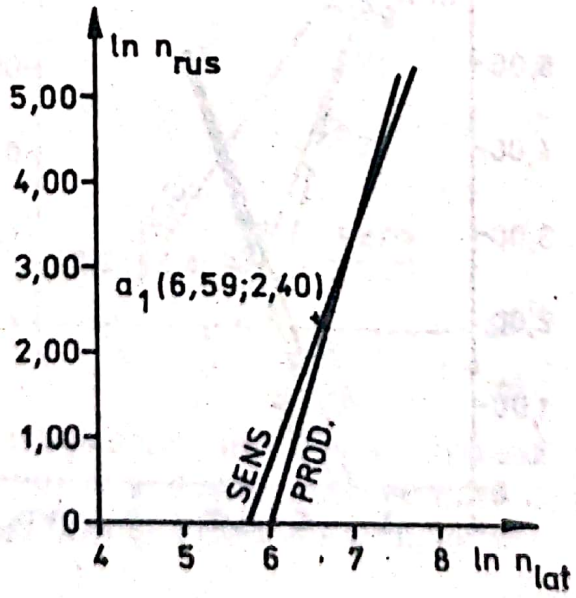
FIG. 19.15




care nu furnizează fiecare decît cîte un singur punct notat cu $a_1$ în figurile 19.14, 19.15 şi 19.16. Demn de semnalat este faptul că, în toate cele trei cazuri menţionate, punctul în cauză cade pe, sau foarte aproape de, dreapta SENS. Există aşadar o concordanţă între gradul de polisemie a cuvintelor de aceste provenienţe şi măsura participării lor la expresii.

Ne oprim aici, reţinînd drept concluzie generală posibilitatea studierii producţiei de expresii şi locuţiuni cu aceleaşi mijloace care ne-au servit la analiza producţiei de derivate şi de sensuri lexicale. Relativa penurie de date de bază, ce s-ar putea fireşte corecta prin extinderea corpusului de expresii considerat, care nu este în nici un caz exhaustiv, ne împiedică să adîncim analogia, prezentînd ilustrări

FIG. 19.16



numerice detaliate ale modelului $M_e$, de tipul celor înfăţişate în tabelul 14.1 şi 18.1 pentru modelele $M_p$ şi, respectiv, *Ms*. Totuşi, chiar dacă, deocamdată, o comparaţie pe întreg frontul componentelor etimologice ale vocabularului nu este posibilă, socotim că una parţială, limitată la acele aporturi pentru care s-au putut construi corelaţii cu fondul latin, prezintă un interes suficient pentru a justifica includerea în lucrarea de faţă a încă unui tabel, pe lîngă atîtea altele. El înfăţişează, ordonate după locul primei ocurenţe lexicale, ponderile numerice ale acestor componente în 22 de secţiuni reprezentative ale vocabularului (tabelul 19.9).

*Tabelul 19.9*

| LATIN | SLAV | BG. & SCR. | BG. | „FR" | TC. | LAT. SAV. | MAGH. | GREC | ONOM. |
|---|---|---|---|---|---|---|---|---|---|
| 10 | — | — | — | — | — | — | — | — | — |
| 13 | 1 | — | — | — | — | — | — | — | — |
| 50 | 4 | — | — | — | — | — | — | — | — |
| 86 | 9 | 1 | — | — | — | — | — | — | — |
| 98 | 11 | 1 | 1 | — | — | — | — | — | — |
| 100 | 12 | 1 | 1 | — | — | — | — | — | — |
| 122 | 16 | 1 | 1 | 1 | — | — | — | — | — |
| 135 | 18 | 1 | 1 | 1 | 1 | — | — | — | — |
| 140 | 19 | 1 | 1 | 1 | 1 | 1 | — | — | — |
| 171 | 26 | 2 | 2 | 2 | 1 | 1 | 1 | — | — |
| 176 | 27 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | — |
| 200 | 33 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | — |
| 255 | 47 | 4 | 4 | 7 | 4 | 3 | 2 | 2 | 1 |
| 400 | 92 | 10 | 12 | 36 | 16 | 9 | 7 | 8 | 3 |
| 600 | 168 | 23 | 30 | 153 | 59 | 26 | 24 | 29 | 12 |
| 800 | 258 | 41 | 41 | 429 | 149 | 58 | 58 | 74 | 37 |
| 1000 | 360 | 64 | 97 | 956 | 304 | 107 | 116 | 153 | 85 |
| 1200 | 472 | 91 | 146 | 1839 | 546 | 176 | 203 | 280 | 168 |
| 1400 | 594 | 123 | 208 | 3198 | 895 | 268 | 326 | 464 | 299 |
| 1600 | 725 | 161 | 283 | 5164 | 1373 | 386 | 491 | 726 | 494 |
| 1800 | 865 | 205 | 371 | 7876 | 2003 | 532 | 705 | 1079 | 768 |
| 2000 | 1012 | 250 | 466 | 11499 | 2807 | 710 | 974 | 1499 | 1141 |



Scanned with OKEN Scanner

În tabelul 19.9, unde, dintre înrîuririle importante, nu lipseşte decît substratul, pot fi urmărite aceleaşi tendinţe generale pe care am avut ocazia să le constatăm deja în cazul productivităţii şi în cel al polisemiei. Structurat acum pe liniile de forţă induse de sistemul expresiilor şi locuţiunilor, vocabularul se dovedeşte a sta tot sub semnul „paradoxului lui A. de Cihac". Ultimele linii ale tabelului ne prezintă, de pildă, elementul turcesc drept superior numericeşte celui latin, ceea ce în mod legic trebuia să survină la un moment dat, în baza regulii generale a răsturnării raportului de forţe dintre oricare două influenţe odată cu creşterea numărului total de cuvinte considerate. Cu mult înainte de a fi depăşit de aportul turcesc, fondul latin cedase întîietatea cantitativă neologismelor „franceze", fenomen evidenţiat deja de modelele $M_p$ şi $M_s$ şi a cărui realitate e uşor de verificat prin simpla comparare a ponderii celor două aporturi într-un lexic „fundamental" de tipul „fond principal" sau „vocabular reprezentativ", în care termenii moşteniţi sînt preponderenţi, şi într-un dicţionar al limbii române, unde aceştia apar striviţi sub masa enormă a împrumuturilor romanice.

Semnificativă este şi ordinea intrării în scenă a influenţelor lexicale. Pentru cele 10 componente reţinute, succesiunea de apariţie în cele trei modele elaborate e cea dată în tabelul 19.10.

*Tabelul 19.10.*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $M_p$ | latin | slav | onom. | bg. & scr. | bg. | magh. | grec | tc. | lat. sav. | „fr" |
| $M_s$ | latin | slav | „fr" | lat. sav. | bg. & scr. | bg. | magh. | tc. | grec | onom. |
| $M_e$ | latin | slav | bg. & scr. | bg. | „fr" | tc. | lat. sav. | magh. | grec | onom. |

Prezenţa componentelor latină şi slavă (întotdeauna în această ordine) pe primele două locuri ale tuturor celor trei clasamente defineşte fără echivoc structura nucleului lexical al limbii române, nucleu alcătuit în exclusivitate din elemente moştenite[1] şi din împrumuturi vechi slave.

[1] Contribuţia substratului nu este pusă în evidenţă în tabelele 19.9 şi 19.10 din pricina dificultăţilor întîmpinate în liniarizarea variaţiei elementului autohton cu fondul latin în cazul modelului $M_e$, aspect discutat pe larg mai înainte. Ea există însă şi nu trebuie în nici un caz neglijată, dacă dorim să obţinem o imagine corectă a constituenţilor sîmburelui vital al vocabularului limbii române.



Scanned with OKEN Scanner

În continuare, ierarhia influenţelor variază de la un model la altul, unele componente etimologice înregistrînd schimbări de poziţie de-a dreptul spectaculoase. Cauzele discrepanţelor constatate între $M_p$ şi $M_s$ au fost deja analizate. Ele rezidă, în esenţă, în faptul că numărul de derivate ale cuvintelor depinde de vîrsta lor, pe cînd numărul de sensuri — nu, afirmaţie valabilă nu în orice condiţii, ci numai pentru o anumită accepţiune a noţiunilor de productivitate şi polisemie. Ne referim la conceptul restrîns de familie de cuvinte, limitat la termenii derivaţi *în româneşte* din unitatea-matcă şi, dimpotrivă, la conceptul extins de polisemie, cuprinzînd nu numai înţelesuri apărute pe teren românesc, ci şi pe cele împrumutate odată cu cuvîntul, ori calchiate ulterior după un model din idiomul sursă sau chiar dintr-o terţă limbă.

Ceea ce pare să fie specific pentru $M_e$ este statutul intermediar, datorat caracterului impur al realităţii modelate: numărul de expresii şi locuţiuni axate pe un anumit cuvînt prezintă asemănări de comportament atît cu numărul de derivate, cît şi cu numărul de înţelesuri ale cuvîntului. Asemenea celui dintîi, numărul de expresii creşte cu trecerea timpului (cuvintele îşi sporesc treptat zestrea de expresii), după cum, urmînd exemplul sensurilor, şi în strînsă legătură cu acestea, el poate înregistra de la bun început valori ridicate, graţie imigrării unor expresii şi locuţiuni asociate deja cuvîntului în limba de origine, fenomen ce poate însoţi sau urma, cu un decalaj temporal mai mare sau mai mic, faptul pătrunderii efective a cuvîntului în limba română. Aşa se face că, de exemplu, elementul „francez“ ocupă, din punctul de vedere al bogăţiei de expresii, o poziţie mai bună decît sub raportul productivităţii, dar nu atît de înaltă ca cea asigurată de criteriul gradului de polisemie. După cum era de aşteptat, situaţia aportului latin savant e perfect analoagă cu a celui „francez“.

Dacă notăm prin $r_p$, $r_s$ şi $r_e$ rangul unei componente etimologice în ierarhia productivităţii şi, respectiv, în cea a polisemiei şi a participării la expresii, ne aşteptăm atunci ca în mai toate cazurile să fie respectate fie dubla inegalitate $r_p \leq r_e \leq r_s$, fie duala ei $r_p \geq r_e \geq r_s$. Nu este însă întotdeauna aşa, regula cunoscînd trei excepţii: aporturile bulgăresc şi turcesc sînt relativ mai bine situate în clasamentul după numărul de expresii şi locuţiuni decît în celelalte două, pe cînd elementul maghiar ocupă, dimpotrivă, o poziţie inferioară din acelaşi punct de vedere celei la care ar fi avut dreptul să aspire în virtutea productivităţii şi polisemiei unităţilor ce-l compun. Un anumit pitoresc al termenilor balcanici poate explica situaţia dintîi. Cealaltă trebuie pusă, probabil, în legătură cu potenţialul mai redus al regionalismelor de a genera locuţiuni şi expresii de mare circulaţie (DEL înregistrează, cum e şi normal pentru o lucrare de dimensiuni mai degrabă modeste, cu precădere formulele lexicale cunoscute tuturor românilor).



Scanned with OKEN Scanner

## 20. O propunere de clasificare multicriterială a componentelor etimologice ale vocabularului

Ne aflăm într-un punct unde, ca o consecinţă logică a demersului analitic precedent, ar fi de aşteptat să urmeze etapa sintezei, integrarea într-un tablou unic a acelor imagini parţiale pe care ni le-a pus la dispoziţie studierea separată a trei importante aspecte ale lexicului, ce constituie totodată trei criterii de evaluare a forţei cuvintelor, şi anume capacitatea de a da naştere la derivate, la sensuri şi la expresii. Căci cum altfel am putea obţine o descriere unitară a ceea ce, cu un termen care poate părea prezumţios, dar nu este decît necesar şi exact, am numit *personalitatea limbii române,* dacă nu încercăm să prindem fizionomia vocabularului într-o formulă integratoare, aptă să înglobeze şi să transceadă în acelaşi timp perspectivele unilaterale sub semnul cărora a stat cercetarea noastră pînă în acest moment? Apelînd la intuiţie, instrument despre care nu vom şti, probabil, niciodată cu certitudine dacă este superior sau inferior cunoaşterii raţionale, Al. Graur a ponderat într-un mod personal, neexplicitabil în termeni algoritmici, cele trei criterii în discuţie, oferindu-ne o selecţie a celor mai importante cuvinte româneşti (fondul principal lexical), întreprindere deopotrivă convingătoare şi discutabilă, în măsura în care cuprinde, pe lîngă un număr mare de rezolvări a căror justeţe e în afara oricărei îndoieli, şi o sumă de opţiuni subiective, ce cu greu ar putea întruni adeziunea tuturor lexicologilor. În fond, era inevitabil ca selecţiei operate de Al. Graur să i se aducă obiecţii, odată ce ea se prezenta ca o alegere personală, care îndreptăţea, prin însuşi precedentul creat, apariţia altor, oricît de multe, încercări similare. Faptul că textul justificativ ce însoţeşte lista fondului principal propus de Al. Graur nu conţine formulări îndeajuns de riguroase ale criteriilor de selecţie pentru a permite delimitarea strictă a cuvintelor aparţinînd FPL de cele rămase în afara acestuia a deschis drum larg feluritelor contestări şi revizuiri, declanşînd un proces căruia nu i se poate pune capăt decît prin reconsiderarea radicală a metodologiei, în sensul supunerii la reguli ferme, clar definite, a căror aplicare să furnizeze un acelaşi rezultat, indiferent cine le aplică. Pe această linie am încercat să mergem în întocmirea celor trei clasificări ale cuvintelor româneşti, din analiza şi generalizarea cărora ne-am extras cam toată substanţa de pînă aici a lucrării.



Scanned with OKEN Scanner

Ideal ar fi acum ca, pornind de la cele trei ierarhii disponibile, să construim o ierarhie de sinteză, constînd în ordonarea neambiguă după importanţă a tuturor cuvintelor limbii române. Proiect ambiţios, ce întîmpină însă, pe lîngă dificultăţi practice lesne de înţeles, şi unele piedici mai subtile, de ordin teoretic. În termenii unei anumite, foarte actuale, matematici a opţiunilor sociale, obiectivul urmărit de noi se subsumează sferei mai largi a problemelor de „decizie multicriterială" şi anume domeniului cunoscut sub denumirea de „agregare a clasamentelor". Or, principalul rezultat obţinut pînă în prezent în studiul rezolubilităţii problemelor de agregare a clasamentelor îl constituie demonstrarea existenţei unor limite inerente acestui tip de probleme. Fără a intra în amănunte (pentru o iniţiere deopotrivă sintetică şi exactă în domeniu, recomandăm consultarea unui opuscul pe cît de accesibil pe atît de atrăgător: *Paradoxurile clasamentelor* de Gh. Păun, Editura ştiinţifică şi enciclopedică, 1987, colecţia „Ştiinţa pentru toţi", din care am extras şi citatele care urmează), trebuie să reţinem că, potrivit unei teoreme celebre a lui K. Arrow „există totdeauna o slăbiciune, o doză de «neraţionalitate», în modul în care se trece de la mai multe ierarhii la una singură, de sinteză". Cum bine observă Gh. Păun, „simplul fapt că există, probabil, mai mult de o sută de metode de agregare a ierarhiilor este un indiciu că «ceva este putred în Danemarca». Într-adevăr, nu poate să nu ne pună pe gînduri constatarea că „dincolo de oricît de multe metode a mai rămas loc pentru una şi încă una şi încă multe altele".

În ceea ce priveşte vocabularul, sîntem, pe de altă parte, în drept să ne întrebăm dacă efortul elaborării unei clasificări unice a cuvintelor, întreprindere a cărei dificultate e uşor de bănuit (să menţionăm doar faptul că, în acest caz, nici nu avem de a face cu o mulţime de obiecte corect definită, avînd în vedere imposibilitatea de a se enumera, element cu element, toţi termenii ce aparţin, la un moment dat, limbii române şi totodată absenţa criteriilor care să permită o formulare analitică a condiţiei de apartenenţă la mulţimea respectivă[1]) se justifică, măcar şi pentru faptul că însemnătatea

[1] Se ştie că, pentru ca o mulţime să fie corect definită, nu e strict necesar ca alcătuirea ei să poată fi prezentată exhaustiv (metodă oricum inoperantă cînd e vorba de mulţimi infinite), fiind suficientă formularea unei/unor condiţii cu ajutorul căreia/cărora să putem decide fără echivoc dacă un obiect dat îi aparţine sau nu. Astfel, mulţimea numerelor naturale pare e perfect definită de clauza divizibilităţii cu 2, care face inutilă enumerarea, de altfel imposibilă, a tuturor elementelor ce o compun.



Scanned with OKEN Scanner

cuvintelor e departe de a se reduce la cei numai trei parametri consideraţi aici.

Cu mult mai interesantă decît instituirea unei ierarhii absolute de importanţă a cuvintelor, proiect ce ar contraveni, prin însăşi rigiditatea premisei sale, caracterului suplu, proteic, pururi schimbător, al limbii, ni se pare încercarea de a evalua, printr-o agregare la alt nivel de generalitate decît cel al unităţilor lexicale individuale, ponderea reală a fiecăreia dintre influenţele exercitate asupra vocabularului în decursul îndelungatei sale istorii. Capitolele precedente ne-au arătat în ce fel pot fi construite acele ecuaţii, reprezentabile sub forma simplă a unor drepte, care să reflecte dinamica raporturilor numerice dintre fiecare aport lexical alogen şi fondul primordial latin al limbii, privite din perspectiva unui anumit criteriu lingvistic. Ca urmare, pentru oricare dintre componentele etimologice ale vocabularului (sau măcar pentru cele principale), am ajuns să dispunem de cîte o tripletă de ecuaţii, respectiv de drepte, corespunzătoare celor trei criterii considerate (productivitate, polisemie, participare la expresii). Ceea ce ne-ar mai rămîne de făcut ar fi ca, pe baza celor trei relaţii cunoscute, să elaborăm o a patra, care să ne ofere expresia integrată, globală, a poziţiei ocupate în limbă de componenta în cauză.

Enunţată astfel, problema nu este însă, din păcate, susceptibilă de o unică rezolvare. Nu-i greu de intuit că se pot imagina un număr practic nelimitat de moduri de mediere, simplă sau ponderată cu diverşi coeficienţi, între valorile furnizate de cele trei ecuaţii, ceea ce va conduce de fiecare dată la un rezultat final diferit.

Pentru a evita impasul, va trebui să explorăm acele posibilităţi de simplificare a problemei care să nu implice sacrificarea unor aspecte esenţiale ale realităţii cercetate. Astfel, dacă ne propunem să renunţăm la una dintre cele trei ecuaţii, apare evident că singura care poate fi, la rigoare, trecută cu vederea este cea care descrie comportarea cuvintelor sub aspectul participării la expresii şi locuţiuni, dată fiind dependenţa lesne sesizabilă a parametrului *număr de expresii generate* de parametrul *număr de sensuri distincte* (cf. începutul capitolului precedent). Păstrînd numai ecuaţiile care reflectă fenomenele independente ale polisemiei şi productivităţii lexicale, pasul următor va trebui să îl constituie precizarea alurei ecuaţiei de sinteză căutate. Desigur, liniaritatea fiecăreia dintre cele două relaţii implicate nu constituie un motiv îndestulător pentru ca şi rezultanta lor să aibă acelaşi caracter. În schimb, variaţia liniară cu factorul latin a diferitelor aporturi alogene în cele trei „cercuri" ale fondului principal lexical stabilit de Al. Graur reprezintă un atare argument, deoarece la autorul selecţiei respective a funcţionat o subiacentă agregare a criteriilor şi odată ce aceasta a condus la o relaţie totalizantă de tip liniar, este de

aşteptat ca şi compunerea explicită a funcţiilor respective să furnizeze un rezultat asemănător.

Convenind asupra liniarităţii rezultantei, să mai observăm că traseul acesteia trebuie să treacă printr-un punct obligat şi anume prin punctul de intersecţie a celor două drepte. Într-adevăr, acest punct corespunde acelei secţiuni prin vocabular în care raportul numeric dintre cele două aporturi lexicale considerate este acelaşi, indiferent dacă ne referim la clasificarea după numărul de derivate sau la cea după numărul de înţelesuri. Astfel, de exemplu, potrivit rezultatelor furnizate de modelele $M_p$ şi $M_s$, în lista celor mai productive 4 571 de cuvinte româneşti un număr de 1209 sînt de origine latină, în timp ce 583 provin din slava veche. Exact tot 1 209 unităţi moştenite din latineşte şi 583 de împrumuturi slave întîlnim şi printre primii 5 608 termeni din lista celor mai „polisemice" cuvinte româneşti. Dacă şi într-un caz şi în celălalt raportul numeric dintre componenta latină şi cea slavă este acelaşi, nu avem nici un motiv să presupunem că în clasamentul „agregat" lucrurile ar sta altminteri. La 1 209 termeni latini moşteniţi e normal să corespundă tot 583 de cuvinte de obîrşie veche slavă, ceea ce înseamnă că punctul de coordonate $x = ln\ 1209 = 7{,}097$, $y = ln\ 583 = 6{,}369$ va fi situat şi pe dreapta asociată relaţiei de sinteză. Concurentă cu liniile corespunzătoare clasificării după productivitate şi, respectiv, după polisemie, dreapta căutată se va situa în unghiul ascuţit format de acestea două, avînd în vedere că, la abscise egale, ea trebuie să prezinte valori intermediare ale ordonatei. În aceste condiţii, ipoteza de mediere cea mai naturală ar fi să considerăm drept rezultantă chiar bisectoarea unghiului format de dreptele date. Ecuaţiile acestora fiind de forma:

$$ln\ n_i = m_1\ ln\ n_{lat} - c_1 \quad \text{şi} \quad ln\ n_i = m_2\ ln\ n_{lat} - c_2$$

se arată uşor (nu mai încărcăm textul cu demonstraţia banală, de geometrie analitică liceală a acestei aserţiuni) că în expresia dreptei rezultante:

$$ln\ n_i = m ln\ n_{lat} - c$$

panta $m$ şi, respectiv, termenul liber $c$ se calculează ca funcţii de parametrii $m_1$, $m_2$, $c_1$, $c_2$ cu ajutorul formulelor:

$$m = \frac{m_1 m_2 - 1 + (1 + m_1^2)(1 + m_2^2)}{m_1 + m_2} \quad \text{şi}$$

$$c = \frac{(m - m_1)\, c_2 - (m - m_2)\, c_1}{m_1 - m_2}$$

Cu ipotezele de mai sus, relaţiile alcătuitoare ale modelelor $M_p$ şi $M_s$ se compun două cîte două, generînd un nou set de 19 ecuaţii ce



Scanned with OKEN Scanner

constituie modelul „agregat" $M_a$. Forma completă a acestuia este următoarea:

$$n_a = exp(1{,}044\ ln\ n_{lat} - 3{,}408)$$
$$n_{sl} = exp(1{,}581\ ln\ n_{lat} - 4{,}851)$$
$$n_{bg.scr} = exp(1{,}521\ ln\ n_{lat} - 6{,}320)$$
$$n_{inc} = exp(1{,}728\ ln\ n_{lat} - 7{,}263)$$
$$n_{bg} = exp(1{,}780\ ln\ n_{lat} - 7{,}743)$$
$$n_{magh} = exp(1{,}773\ ln\ n_{lat} - 7{,}857)$$
$$n_{onom} = exp(1{,}981\ ln\ n_{lat} - 9{,}189)$$
$$n_{nec} = exp(2{,}222\ ln\ n_{lat} - 10{,}287)$$
$$n_{fr} = exp(3{,}186\ ln\ n_{lat} - 15{,}200)$$
$$n_{lat.sav} = exp(2{,}344\ ln\ n_{lat} - 11{,}469)$$
$$n_{gr} = exp(2{,}299\ ln\ n_{lat} - 11{,}570)$$
$$n_{germ} = exp(2{,}663\ ln\ n_{lat} - 14{,}637)$$
$$n_{n.s.m.} = exp(1{,}893\ ln\ n_{lat} - 10{,}069)$$
$$n_{tc} = exp(2{,}738\ ln\ n_{lat} - 14{,}264)$$
$$n_{scr} = exp(2{,}398\ ln\ n_{lat} - 13{,}267)$$
$$n_{rom(lat)} = exp(2{,}494\ ln\ n_{lat} - 13{,}947)$$
$$n_{it} = exp(3{,}279\ ln\ n_{lat} - 19{,}083)$$
$$n_{rus} = exp(3{,}073\ ln\ n_{lat} - 18{,}074)$$
$$n_{ucr} = exp(3{,}171\ ln\ n_{lat} - 18{,}699)$$

Structural identic cu $M_p$ şi $M_s$, modelul $M_a$ diferă de acestea numai prin valorile coeficienţilor $m$ şi $c$, care, după cum s-a arătat, sînt calculaţi în aşa fel încît să conducă în toate cazurile la rezultate intermediare în raport cu cele furnizate de modelele productivităţii şi polisemiei. Că lucrurile stau într-adevăr aşa se poate verifica şi comparînd valorile numerice din tabelele 14.4 şi 18.1 cu omoloagele lor din tabelul 20, în care sînt sistematizate, după un tipic deja cunoscut cititorului, rezultatele aplicării modelului $M_a$.

În legătură cu această comparaţie, e strict necesar să subliniem deosebirea esenţială dintre ultimul tabel şi precedentele. Dacă acestea reflectau anumite situaţii concrete (pe alocuri uşor ajustate prin modelare, dar oricum solid ancorate în realitatea lingvistică), tabelul 20, ca şi modelul $M_a$ pe care îl ilustrează, se prezintă, în schimb, drept una dintre acele abstracţiuni pe care un spirit prea exigent le-ar putea socoti „fără acoperire".



Scanned with OKEN Scanner

Într-adevăr, valorile absolute ale datelor din tabelul 20 nu pot fi coroborate printr-o exemplificare lingvistică directă, de felul celei posibile în cazul tabelelor 14.1 şi 18.1. Astfel, dacă, din calcul, a rezultat că printre cele mai productive 100 de cuvinte româneşti un număr de 13 trebuie să fi provenit din slava veche, aceste 13 unităţi prognozate pot fi şi efectiv nominalizate. E vorba de cuvintele *clipi, drag, gol, lipi, pod, potrivă, prost, rob, suci, tîrî, ţigan, ţintă, val.* La fel, în ierarhia polisemiei, conform ecuaţiilor modelului $M_s$, printre primele 1012 cuvinte e de aşteptat să existe 16 termeni de origine maghiară, previziune confirmată de excerptarea Dicţionarului Limbii Române Moderne, unde presupusele 16 elemente pot fi realmente identificate. Ele sînt: *bănat, bănui, chip, corlată, gînd, majă, meşteşug, neam, pecetlui, pont, sălaş, seamă, socoti, tobă, ţarcă, zăbală.* Nu acelaşi lucru se poate spune despre modelul $M_a$ şi tabelul 20. Conform lor, în rîndul celor mai importante 500 de cuvinte din limba română e probabil să întîlnim 22 de neologisme de sorginte latină savantă, cărora însă, de data aceasta, nu sîntem în măsură să le precizăm identitatea.

Să deducem de aici lipsa de valoare practică a modelului „agregat"? Cîtuşi de puţin, căci, în ciuda „slăbiciunii" menţionate, acesta se dovedeşte singurul capabil să ne ofere acea imagine echilibrată a raporturilor dintre componentele etimologice ale limbii care caracterizează, dincolo de fluctuaţiile unuia sau altuia dintre parametrii lexicali, un anumit „regim mediu" al vocabularului.

Ştim, bunăoară, că însemnătatea unor înrîuriri lexicale asupra limbii române este semnificativ caracterizată de adîncimea de pătrundere a împrumuturilor respective în straturile profunde ale vocabularului. Din acest punct de vedere, ordinea apariţiei diferitelor aporturi în tabelele de tipul 14.1, 18.1 şi 20 se constituie într-o veritabilă ierarhie de importanţă a componentelor etimologice ale limbii. În legătură cu aportul „francez", am văzut că, privit din unghiul productivităţii, el se situează, ca adîncime de penetraţie în nucleul vital al limbii, pe un modest loc 15, în timp ce din unghi semantic deţine, dimpotrivă, o poziţie fruntaşă, a treia. Or, e de la sine înţeles că însemnătatea globală a neologismelor „franceze" nu este nici atît de mare pe cît ne-o relevă studiul polisemiei, dar nici atît de redusă cum rezultă din analiza potenţialului lor lexicogen. Calea de mijloc întruchipată de modelul $M_a$ pare să definească mult mai exact locul influenţei „franceze" asupra limbii române, în ciuda faptului că valorilor particulare calculate cu ajutorul acestui model nu le putem asocia liste concrete de unităţi lexicale, similare celor ce ilustrează modelele $M_p$ şi $M_s$. Consultînd tabelul 20, constatăm, de exemplu, că ponderea elementului „francez" în ansamblul celor mai importante 300 de cuvinte româneşti ar trebui să fie de cca. 2% (6 unităţi). Care sînt însă cei 6 exponenţi ai acestui



Scanned with OKEN Scanner

*Tabelul 20*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | BG. & SCR. | INC. | BG. | MAGH. | ONOM. | NEC. | „FR." | LAT. SAV. | GR. | NSM | TC. | GERM. | SCR. | ROM. (LAT) | IT. | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 10 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 16 | 14 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 20 | 18 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 30 | 27 | 1 | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 40 | 37 | 1 | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 47 | 41 | 2 | 3 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 50 | 44 | 2 | 3 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 52 | 45 | 2 | 3 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 60 | 52 | 2 | 4 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 62 | 53 | 2 | 4 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 68 | 57 | 2 | 5 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 70 | 59 | 2 | 5 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 80 | 67 | 3 | 6 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 88 | 73 | 3 | 7 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 90 | 75 | 3 | 7 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — | — |
| 92 | 76 | 3 | 7 | 1 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 100 | 83 | 3 | 8 | 1 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — | — |
| 118 | 95 | 4 | 10 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 120 | 96 | 4 | 11 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — | — |
| 126 | 100 | 4 | 11 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 140 | 111 | 4 | 13 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — | — |
| 147 | 114 | 5 | 14 | 2 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — | — | — | — | — | — |
| 160 | 123 | 5 | 16 | 3 | 3 | 2 | 2 | 1 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — | — |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 180 | 137 | 6 | 18 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | — | — | — | — | — | — | — | — |
| **190** | 142 | 6 | 20 | 3 | 4 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | — | — | — | — | — | — | — |
| **192** | 143 | 6 | 20 | 3 | 4 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — |
| 200 | 149 | 6 | 21 | 4 | 4 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — | — | — |
| 220 | 160 | 7 | 24 | 4 | 5 | 4 | 3 | 2 | 3 | 3 | 2 | 1 | 1 | 1 | — | — | — | — | — | — |
| 240 | 174 | 7 | 27 | 5 | 5 | 4 | 4 | 3 | 3 | 3 | 2 | 1 | 1 | 1 | — | — | — | — | — | — |
| 260 | 185 | 8 | 30 | 5 | 6 | 5 | 4 | 3 | 4 | 4 | 2 | 2 | 1 | 1 | — | — | — | — | — | — |
| **265** | 188 | 8 | 31 | 5 | 6 | 5 | 4 | 3 | 4 | 4 | 2 | 2 | 1 | 1 | 1 | — | — | — | — | — |
| **269** | 190 | 8 | 31 | 5 | 6 | 5 | 4 | 3 | 4 | 5 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — |
| 280 | 197 | 8 | 33 | 6 | 6 | 5 | 4 | 4 | 4 | 5 | 2 | 2 | 1 | 1 | 1 | 1 | — | — | — | — |
| **297** | 204 | 9 | 35 | 6 | 7 | 6 | 5 | 4 | 5 | 6 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — |
| 300 | 205 | 9 | 35 | 6 | 7 | 6 | 5 | 4 | 5 | 6 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | — | — | — |
| 320 | 219 | 9 | 39 | 7 | 8 | 6 | 5 | 4 | 5 | 7 | 3 | 2 | 1 | 2 | 1 | 1 | 1 | — | — | — |
| 340 | 228 | 10 | 42 | 7 | 8 | 7 | 6 | 5 | 6 | 8 | 4 | 3 | 1 | 2 | 1 | 1 | 1 | — | — | — |
| 360 | 239 | 10 | 45 | 8 | 9 | 8 | 6 | 5 | 7 | 10 | 4 | 4 | 1 | 2 | 1 | 1 | 1 | — | — | — |
| 380 | 251 | 11 | 48 | 8 | 10 | 8 | 7 | 6 | 7 | 11 | 4 | 3 | 1 | 2 | 1 | 1 | 1 | — | — | — |
| 400 | 261 | 11 | 52 | 8 | 10 | 9 | 7 | 6 | 8 | 12 | 5 | 3 | 2 | 3 | 1 | 1 | 1 | — | — | — |
| 420 | 270 | 11 | 55 | 9 | 11 | 9 | 8 | 7 | 9 | 14 | 5 | 4 | 2 | 3 | 1 | 1 | 1 | — | — | — |
| **427** | 273 | 12 | 56 | 9 | 11 | 9 | 8 | 7 | 9 | 14 | 5 | 4 | 2 | 3 | 1 | 1 | 1 | 1 | — | — |
| 440 | 280 | 12 | 58 | 9 | 12 | 10 | 8 | 7 | 9 | 16 | 6 | 4 | 2 | 3 | 1 | 1 | 1 | 1 | — | — |
| **454** | 286 | 12 | 60 | 10 | 12 | 10 | 9 | 8 | 10 | 17 | 6 | 4 | 2 | 3 | 2 | 1 | 1 | 1 | 1 | — |
| 460 | 288 | 12 | 61 | 10 | 13 | 11 | 9 | 8 | 10 | 17 | 6 | 4 | 2 | 3 | 2 | 1 | 1 | 1 | 1 | — |
| **469** | 293 | 12 | 62 | 10 | 13 | 11 | 9 | 8 | 10 | 18 | 6 | 4 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | 1 |
| 480 | 297 | 13 | 64 | 10 | 13 | 11 | 9 | 8 | 11 | 19 | 7 | 5 | 2 | 4 | 2 | 1 | 1 | 1 | 1 | 1 |
| 500 | 306 | 13 | 67 | 11 | 14 | 12 | 10 | 9 | 11 | 21 | 7 | 5 | 2 | 4 | 2 | 2 | 1 | 1 | 1 | 1 |
| 520 | 317 | 13 | 70 | 11 | 15 | 12 | 10 | 9 | 12 | 23 | 8 | 5 | 2 | 4 | 2 | 2 | 2 | 1 | 1 | 1 |



Scanned with OKEN Scanner

*Tabelul 20 (continuare)*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | BG. & SCR. | INC. | BG. | MAGH. | ONOM. | NEC. | „FR." | LAT. SAV. | GR. | NSM | TC. | GERM. | SCR. | ROM. (LAT) | IT. | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 540 | 324 | 14 | 73 | 12 | 15 | 13 | 11 | 10 | 13 | 25 | 8 | 6 | 2 | 5 | 2 | 2 | 2 | 1 | 1 | 1 |
| 560 | 333 | 14 | 76 | 12 | 16 | 14 | 12 | 10 | 14 | 27 | 9 | 6 | 3 | 5 | 2 | 2 | 2 | 1 | 1 | 1 |
| 580 | 341 | 15 | 79 | 13 | 17 | 14 | 12 | 11 | 15 | 30 | 9 | 6 | 3 | 6 | 2 | 2 | 2 | 1 | 1 | 1 |
| 600 | 351 | 15 | 82 | 13 | 17 | 15 | 13 | 11 | 15 | 32 | 10 | 7 | 3 | 6 | 3 | 2 | 2 | 1 | 1 | 1 |
| 620 | 360 | 15 | 86 | 14 | 18 | 15 | 13 | 12 | 16 | 35 | 10 | 7 | 3 | 6 | 3 | 2 | 2 | 1 | 1 | 1 |
| 640 | 368 | 16 | 89 | 14 | 19 | 16 | 14 | 12 | 17 | 37 | 11 | 7 | 3 | 7 | 3 | 2 | 2 | 1 | 1 | 1 |
| 660 | 375 | 16 | 92 | 15 | 20 | 17 | 14 | 13 | 18 | 40 | 11 | 8 | 3 | 7 | 3 | 3 | 2 | 1 | 1 | 1 |
| 680 | 384 | 16 | 95 | 15 | 20 | 17 | 15 | 13 | 19 | 43 | 12 | 8 | 3 | 8 | 3 | 3 | 2 | 2 | 1 | 1 |
| 700 | 390 | 17 | 98 | 16 | 21 | 18 | 15 | 14 | 20 | 45 | 12 | 9 | 3 | 8 | 4 | 3 | 3 | 2 | 1 | 1 |
| 720 | 399 | 17 | 101 | 16 | 22 | 19 | 16 | 14 | 20 | 48 | 13 | 9 | 4 | 8 | 4 | 3 | 3 | 2 | 1 | 1 |
| 740 | 406 | 18 | 104 | 17 | 23 | 19 | 16 | 15 | 21 | 51 | 14 | 9 | 4 | 9 | 4 | 3 | 3 | 2 | 1 | 1 |
| 760 | 414 | 18 | 107 | 17 | 23 | 20 | 17 | 16 | 22 | 54 | 14 | 10 | 4 | 9 | 4 | 3 | 3 | 2 | 2 | 1. |
| 780 | 420 | 18 | 110 | 18 | 24 | 21 | 17 | 16 | 23 | 58 | 15 | 10 | 4 | 10 | 4 | 3 | 3 | 2 | 2 | 2 |
| 800 | 428 | 18 | 113 | 18 | 25 | 21 | 18 | 17 | 24 | 61 | 15 | 11 | 4 | 10 | 4 | 4 | 3 | 2 | 2 | 2 |
| 820 | 435 | 19 | 116 | 19 | 25 | 22 | 18 | 17 | 25 | 64 | 16 | 11 | 4 | 11 | 5 | 4 | 3 | 2 | 2 | 2 |
| 840 | 443 | 19 | 119 | 19 | 26 | 22 | 19 | 18 | 26 | 68 | 17 | 11 | 4 | 11 | 5 | 4 | 3 | 2 | 2 | 2 |
| 860 | 449 | 19 | 122 | 19 | 27 | 23 | 20 | 18 | 27 | 71 | 17 | 12 | 4 | 12 | 5 | 4 | 4 | 3 | 2 | 2 |
| 880 | 455 | 20 | 125 | 20 | 28 | 24 | 20 | 19 | 28 | 74 | 18 | 12 | 5 | 12 | 5 | 4 | 4 | 3 | 2 | 2 |
| 900 | 462 | 20 | 128 | 20 | 28 | 24 | 21 | 19 | 29 | 78 | 19 | 13 | 5 | 13 | 6 | 4 | 4 | 3 | 2 | 2 |
| 920 | 471 | 20 | 131 | 21 | 29 | 25 | 21 | 20 | 29 | 82 | 19 | 13 | 5 | 13 | 6 | 4 | 4 | 3 | 2 | 2 |
| 940 | 475 | 21 | 134 | 21 | 30 | 26 | 22 | 21 | 30 | 85 | 20 | 14 | 5 | 14 | 6 | 5 | 4 | 3 | 2 | 2 |
| 960 | 484 | 21 | 137 | 22 | 30 | 26 | 22 | 21 | 31 | 90 | 20 | 14 | 5 | 14 | 6 | 5 | 4 | 3 | 3 | 2 |



Scanned with OKEN Scanner

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 980 | 490 | 21 | 140 | 22 | 31 | 27 | 23 | 22 | 32 | 93 | 21 | 14 | 5 | 15 | 6 | 5 | 4 | 3 | 3 | 3 |
| 1000 | 495 | 22 | 143 | 23 | 32 | 27 | 23 | 22 | 33 | 96 | 22 | 15 | 5 | 15 | 7 | 5 | 5 | 4 | 3 | 3 |
| 1050 | 510 | 22 | 150 | 24 | 34 | 29 | 25 | 24 | 36 | 107 | 23 | 16 | 6 | 17 | 7 | 5 | 5 | 4 | 3 | 3 |
| 1100 | 527 | 23 | 157 | 25 | 35 | 31 | 26 | 25 | 38 | 118 | 25 | 17 | 6 | 18 | 8 | 6 | 5 | 4 | 3 | 3 |
| 1150 | 542 | 24 | 164 | 26 | 37 | 32 | 27 | 27 | 40 | 129 | 27 | 18 | 6 | 20 | 8 | 6 | 6 | 5 | 4 | 4 |
| 1200 | 557 | 24 | 171 | 27 | 39 | 34 | 28 | 28 | 43 | 139 | 28 | 19 | 7 | 21 | 9 | 7 | 6 | 5 | 4 | 4 |
| 1250 | 571 | 25 | 178 | 28 | 41 | 35 | 30 | 29 | 45 | 151 | 30 | 20 | 7 | 22 | 10 | 7 | 7 | 6 | 4 | 4 |
| 1300 | 585 | 26 | 185 | 29 | 42 | 37 | 31 | 31 | 48 | 163 | 32 | 22 | 7 | 24 | 10 | 7 | 7 | 6 | 4 | 4 |
| 1350 | 597 | 26 | 191 | 30 | 44 | 38 | 32 | 32 | 50 | 175 | 34 | 23 | 8 | 26 | 11 | 8 | 7 | 7 | 5 | 5 |
| 1400 | 610 | 27 | 198 | 31 | 46 | 40 | 34 | 34 | 53 | 188 | 35 | 24 | 8 | 27 | 12 | 8 | 8 | 7 | 5 | 5 |
| 1450 | 625 | 27 | 205 | 32 | 47 | 41 | 35 | 35 | 55 | 201 | 37 | 25 | 8 | 29 | 12 | 9 | 8 | 8 | 6 | 5 |
| 1500 | 635 | 28 | 212 | 33 | 49 | 43 | 36 | 37 | 58 | 214 | 39 | 26 | 9 | 30 | 13 | 9 | 9 | 8 | 6 | 6 |
| 1550 | 648 | 29 | 218 | 34 | 51 | 44 | 37 | 38 | 60 | 227 | 41 | 28 | 9 | 32 | 14 | 10 | 9 | 9 | 6 | 6 |
| 1600 | 660 | 29 | 225 | 35 | 52 | 46 | 39 | 39 | 63 | 241 | 43 | 29 | 9 | 34 | 14 | 10 | 9 | 9 | 7 | 7 |
| 1650 | 672 | 30 | 231 | 36 | 54 | 47 | 40 | 41 | 66 | 255 | 44 | 30 | 10 | 35 | 15 | 10 | 10 | 10 | 7 | 7 |
| 1700 | 685 | 30 | 237 | 37 | 56 | 49 | 41 | 42 | 68 | 270 | 46 | 31 | 10 | 37 | 16 | 11 | 10 | 10 | 7 | 7 |
| 1750 | 695 | 31 | 244 | 38 | 57 | 50 | 42 | 44 | 71 | 284 | 48 | 32 | 10 | 39 | 16 | 11 | 11 | 11 | 8 | 8 |
| 1800 | 706 | 31 | 250 | 39 | 59 | 52 | 44 | 45 | 73 | 300 | 50 | 34 | 10 | 40 | 17 | 12 | 11 | 11 | 8 | 8 |
| 1850 | 717 | 32 | 256 | 40 | 60 | 53 | 45 | 46 | 76 | 314 | 52 | 35 | 11 | 42 | 18 | 12 | 12 | 12 | 8 | 9 |
| 1900 | 729 | 32 | 262 | 41 | 62 | 54 | 46 | 48 | 78 | 330 | 53 | 36 | 11 | 44 | 18 | 13 | 12 | 13 | 9 | 9 |
| 1950 | 740 | 33 | 268 | 42 | 63 | 56 | 47 | 49 | 81 | 345 | 55 | 37 | 11 | 46 | 19 | 13 | 13 | 13 | 9 | 9 |
| 2000 | 750 | 33 | 274 | 42 | 65 | 57 | 48 | 51 | 83 | 361 | 57 | 38 | 12 | 48 | 20 | 14 | 13 | 14 | 10 | 10 |
| 2200 | 790 | 35 | 298 | 46 | 71 | 63 | 53 | 56 | 94 | 426 | 65 | 43 | 13 | 55 | 23 | 15 | 15 | 16 | 11 | 12 |
| 2400 | 828 | 37 | 320 | 49 | 77 | 68 | 58 | 62 | 104 | 495 | 72 | 48 | 14 | 62 | 26 | 17 | 17 | 19 | 13 | 14 |
| 2600 | 863 | 38 | 343 | 53 | 83 | 74 | 62 | 67 | 114 | 567 | 80 | 53 | 15 | 70 | 29 | 19 | 18 | 22 | 15 | 15 |
| 2800 | 897 | 40 | 364 | 56 | 89 | 79 | 67 | 72 | 124 | 641 | 87 | 58 | 16 | 78 | 32 | 21 | 20 | 25 | 17 | 17 |



Scanned with OKEN Scanner

*Tabelul 20 (continuare)*

| TOTAL CUVINTE | LAT. | cf. ALB. | SLAVE | BG. & SCR. | INC. | BG. | MAGH. | ONOM. | NEC. | „FR." | LAT. SAV. | GR. | NSM | TC. | GERM. | SCR. | ROM. (LAT) | IT. | RUS | UCR. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3000 | 929 | 42 | 385 | 59 | 94 | 84 | 71 | 77 | 134 | 717 | 95 | 63 | 18 | 85 | 35 | 23 | 22 | 28 | 19 | 20 |
| 3200 | 960 | 43 | 406 | 62 | 100 | 89 | 75 | 83 | 144 | 795 | 102 | 68 | 19 | 93 | 38 | 25 | 24 | 31 | 21 | 22 |
| 3400 | 990 | 44 | 426 | 65 | 105 | 94 | 79 | 88 | 154 | 875 | 110 | 73 | 20 | 102 | 42 | 26 | 26 | 34 | 23 | 24 |
| 3600 | 1018 | 46 | 445 | 68 | 110 | 99 | 83 | 93 | 164 | 958 | 117 | 78 | 21 | 110 | 45 | 28 | 28 | 38 | 25 | 26 |
| 3800 | 1045 | 47 | 465 | 70 | 116 | 104 | 87 | 98 | 174 | 1043 | 125 | 82 | 22 | 118 | 48 | 30 | 30 | 41 | 27 | 28 |
| 4000 | 1072 | 48 | 483 | 73 | 121 | 108 | 91 | 103 | 184 | 1128 | 132 | 87 | 23 | 126 | 52 | 32 | 32 | 45 | 29 | 31 |
| 4200 | 1097 | 49 | 501 | 76 | 126 | 113 | 95 | 108 | 194 | 1215 | 140 | 92 | 24 | 135 | 55 | 34 | 34 | 48 | 31 | 33 |
| 4400 | 1122 | 51 | 519 | 78 | 131 | 118 | 99 | 113 | 204 | 1303 | 147 | 97 | 25 | 143 | 58 | 36 | 35 | 52 | 33 | 36 |
| 4600 | 1146 | 52 | 536 | 81 | 136 | 122 | 103 | 117 | 213 | 1395 | 154 | 102 | 26 | 151 | 62 | 38 | 37 | 55 | 36 | 38 |
| 4800 | 1169 | 53 | 554 | 83 | 140 | 126 | 106 | 122 | 223 | 1488 | 162 | 107 | 27 | 160 | 65 | 39 | 39 | 59 | 38 | 40 |
| 5000 | 1191 | 54 | 570 | 86 | 145 | 131 | 110 | 127 | 233 | 1580 | 169 | 111 | 28 | 169 | 68 | 41 | 41 | 63 | 40 | 43 |
| 6000 | 1295 | 59 | 650 | 98 | 167 | 151 | 128 | 149 | 280 | 2060 | 206 | 135 | 33 | 212 | 85 | 50 | 51 | 83 | 52 | 56 |
| 7000 | 1387 | 63 | 725 | 108 | 188 | 171 | 144 | 171 | 326 | 2564 | 242 | 158 | 38 | 256 | 102 | 59 | 60 | 104 | 64 | 70 |
| 8000 | 1470 | 67 | 795 | 118 | 208 | 190 | 160 | 192 | 371 | 3085 | 278 | 181 | 42 | 299 | 120 | 68 | 70 | 125 | 77 | 84 |
| 9000 | 1546 | 71 | 861 | 128 | 212 | 208 | 175 | 227 | 415 | 3624 | 312 | 203 | 46 | 344 | 137 | 77 | 79 | 148 | 89 | 98 |
| 10000 | 1616 | 74 | 924 | 137 | 245 | 225 | 189 | 232 | 459 | 4175 | 346 | 225 | 50 | 389 | 154 | 86 | 88 | 171 | 102 | 113 |
| 15000 | 1907 | 88 | 1201 | 176 | 327 | 302 | 253 | 322 | 663 | 7086 | 511 | 329 | 69 | 612 | 239 | 127 | 133 | 294 | 170 | 191 |
| 20000 | 2136 | 99 | 1437 | 209 | 397 | 370 | 310 | 403 | 853 | 10163 | 666 | 427 | 85 | 835 | 324 | 167 | 177 | 427 | 241 | 274 |
| 25000 | 2328 | 108 | 1646 | 238 | 461 | 431 | 361 | 478 | 1032 | 13354 | 815 | 520 | 100 | 1057 | 407 | 205 | 219 | 566 | 314 | 360 |
| 30000 | 2494 | 116 | 1835 | 264 | 519 | 487 | 408 | 548 | 1203 | 16634 | 958 | 609 | 114 | 1276 | 489 | 242 | 260 | 709 | 388 | 447 |



Scanned with OKEN Scanner

aport e mai dificil de precizat, avînd în vedere şi faptul că mulţimea celor mai productive împrumuturi „franceze" se dovedeşte total disjunctă de cea a împrumuturilor celor mai polisemice. Astfel, după numărul lor de derivate, neologismele „franceze" cele mai bine situate sînt *alimenta, aprecia, lustru, tramă, valoare, califica, pachet, soldat, şosea,* pe cînd din punctul de vedere al numărului de înţelesuri pe primele locuri se situează *corp, centru, clasă, curent, forţă, nervură, ton.* Luarea în considerare, în plus, a clasificării după numărul de expresii şi locuţiuni nu e deloc de natură să uşureze răspunsul, ci, din contră, complică şi mai mult lucrurile, aducînd în discuţie o nouă colecţie de 9 elemente total diferite de precedentele: *idee, notă, onoare, circulaţie, cont, parolă, raport, rezervă, termen.* Încercarea de a construi totuşi un clasament de importanţă a cuvintelor de origine „franceză" (ghilimelele sînt aici mai la locul lor decît oriunde, căci, termeni ca *termen* şi *centru* provin în mod vădit din latina literară, iar *forţă* — din italiană) este oricînd posibilă, dar trebuie spus că ea ne poartă de-a dreptul în zona periculoasă a agregării ierarhiilor, teren pe care dorim să îl evităm spre a diminua la minimum numărul conjecturilor pasibile de a suscita obiecţii de un fel sau altul.

De fapt, ceea ce ne interesează în mod special la modelul $M_a$ nu este ierarhizarea cuvintelor de diferite provenienţe, ci ierarhizarea provenienţelor înseşi. Potrivit tabelului 20, ordinea aporturilor lexicale după adîncimea lor de pătrundere în miezul vocabularului (recte după rangul în tabel al liniei în cadrul căreia survine ocurenţa primului element de originea respectivă) este următoarea: fondul latin moştenit ($n = 1$), substratul şi elementul vechi slav (ambele cu $n = 15$), materialul lexical cu dublă etimologie bulgaro-sîrbo-croată ($n = 47$), cuvintele de origine incertă ($n = 52$), elementul bulgăresc ($n = 63$), cel maghiar ($n = 68$), formaţiile onomatopeice ($n = 89$), termenii cu etimologie necunoscută ($n = 93$), neologismele „franceze" ($n = 119$), cele latine savante ($n = 126$), aportul grecesc ($n = 147$), cel intitulat „nord-slav medieval" ($n = 190$), cel turcesc ($n = 191$), împrumuturile din germană ($n = 266$), din sîrbocroată ($n = 269$), cuvintele construite în româneşte din elemente de origine latină ($n = 294$), aportul italian ($n = 427$), cel rusesc ($n = 454$) şi cel ucrainian ($n = 469$).

Lăsînd la o parte categoriile eteroclite (cuvintele de provenienţă incertă, necunoscută, mixtă), se conturează o clasificare conform căreia diversele influenţe suferite de limba română de-a lungul timpului pot fi grupate în patru clase de importanţă descrescătoare şi anume:

I. Aporturile vechi: substratul, elementul vechi slav, cel bulgăresc, cel maghiar şi formaţiile onomatopeice.

II. Afluxul neologic „francez" şi latin savant care a reromanizat în ultimii 100—150 de ani vocabularul limbii române.



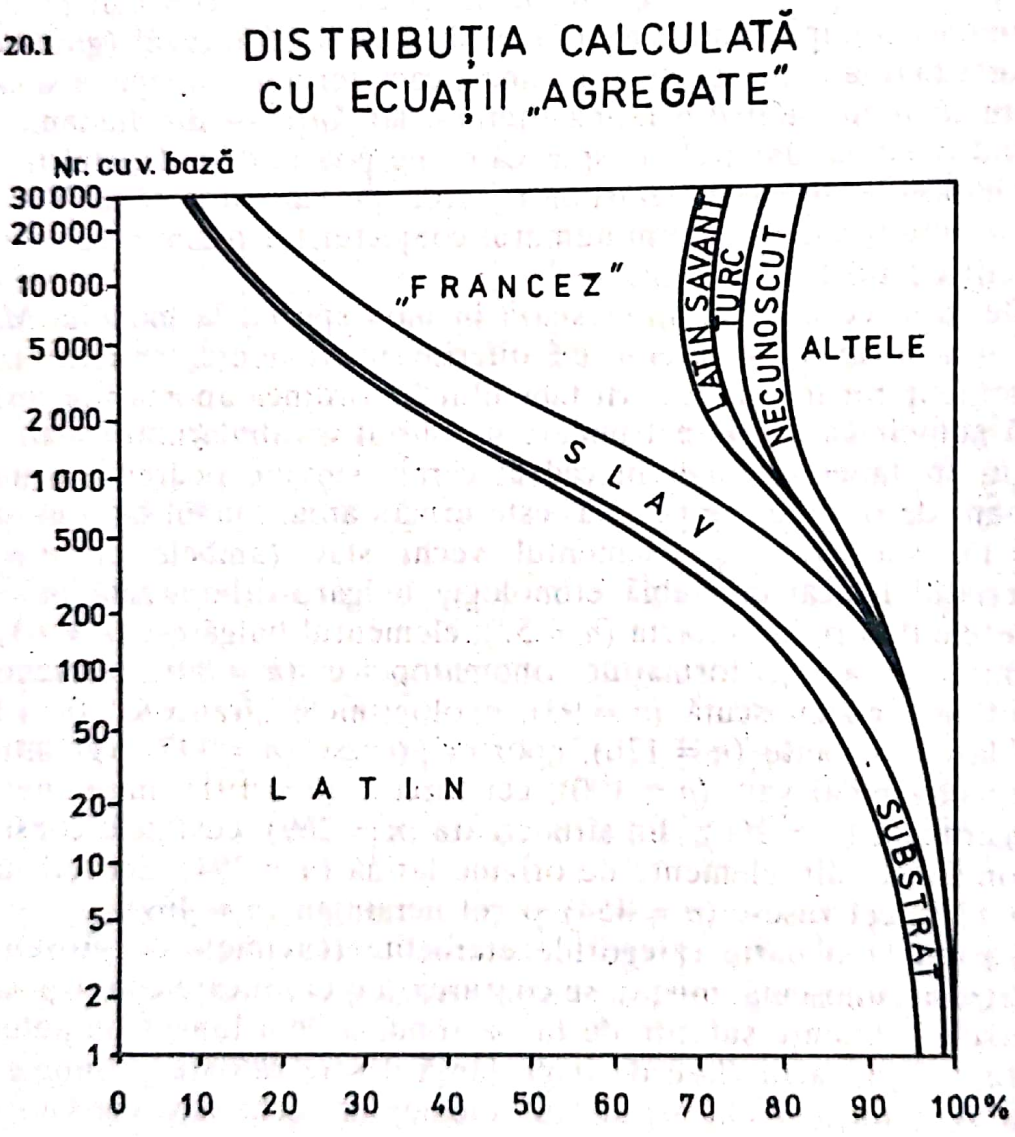
Scanned with OKEN Scanner

III. Înrîuririle premoderne grecească şi turcească, active cu precădere în epoca fanariotă.

IV. Unele influenţe marginale, de însemnătate mai redusă, cum ar fi cea germană, cea sîrbocroată, italiană, rusească, ucraineană etc[1].

Într-o măsură mai mare sau mai mică, toate aceste contribuţii conferă limbii române coloratura ei specifică, esenţa vocabularului rămînînd însă copios dominată de zestrea latină. O imagine sugestivă a acestei predominanţe absolute (nu întotdeauna pusă în evidenţă de statisticile lexicale empirice) ne-o oferă reprezentarea grafică a ponderii procentuale a diferitelor componente etimologice (fig. 20.1).

FIG. 20.1

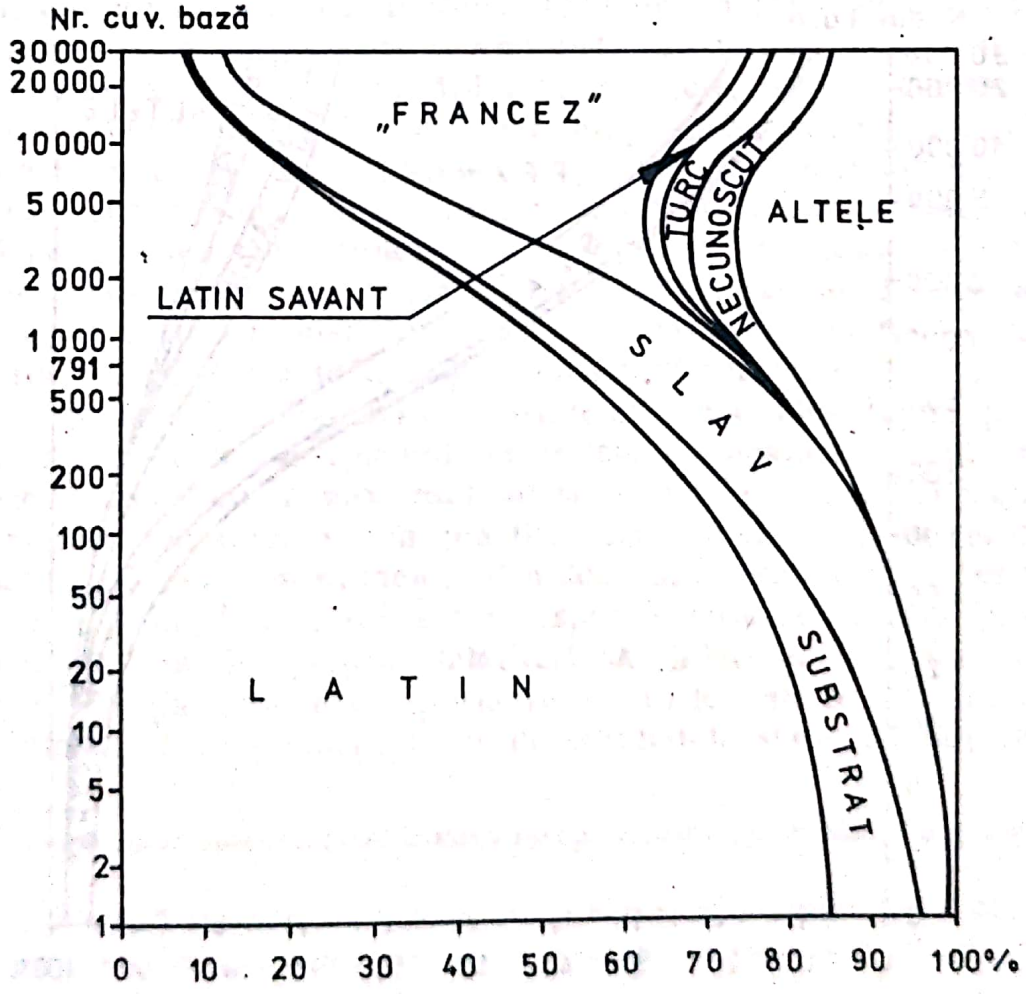



[1] Lista rămîne deschisă, deoarece aici trebuie incluse şi alte aporturi neluate de noi explicit în considerare, ca cel polon, cel ţigănesc etc.

Topograma obţinută constituie cea mai elocventă dovadă a caracterului fundamental latin al lexicului limbii române. Scara logaritmică în care este reprezentat, pe ordonată, numărul de cuvinte-bază conferă, cum e şi firesc, un spaţiu mai mare zonei unităţilor mai importante, adică tocmai domeniului lexical în cuprinsul căruia fondul latin deţine majoritatea absolută. Aşa se face că, deşi la nivel de dicţionar prezenţa latină moştenită e modestă, aria ce-i revine în reprezentarea grafică din figura 20.1 depăşeşte 65% din întreaga suprafaţă a topogramei. Ţinînd seama de faptul că încă vreo 10 procente din suprafaţă sînt ocupate de componentele „franceză" şi latină savantă, urmează că ponderea globală a cuvintelor de sursă nelatină e limitată la mai puţin de un sfert din ansamblul ordonat al vocabularului.

Susceptibile de reprezentări asemănătoare sînt şi rezultatele privitoare la distribuţia pe origini a cuvintelor ordonate după productivitate (modelul $M_p$ — fig. 20.2) şi după numărul de înţelesuri

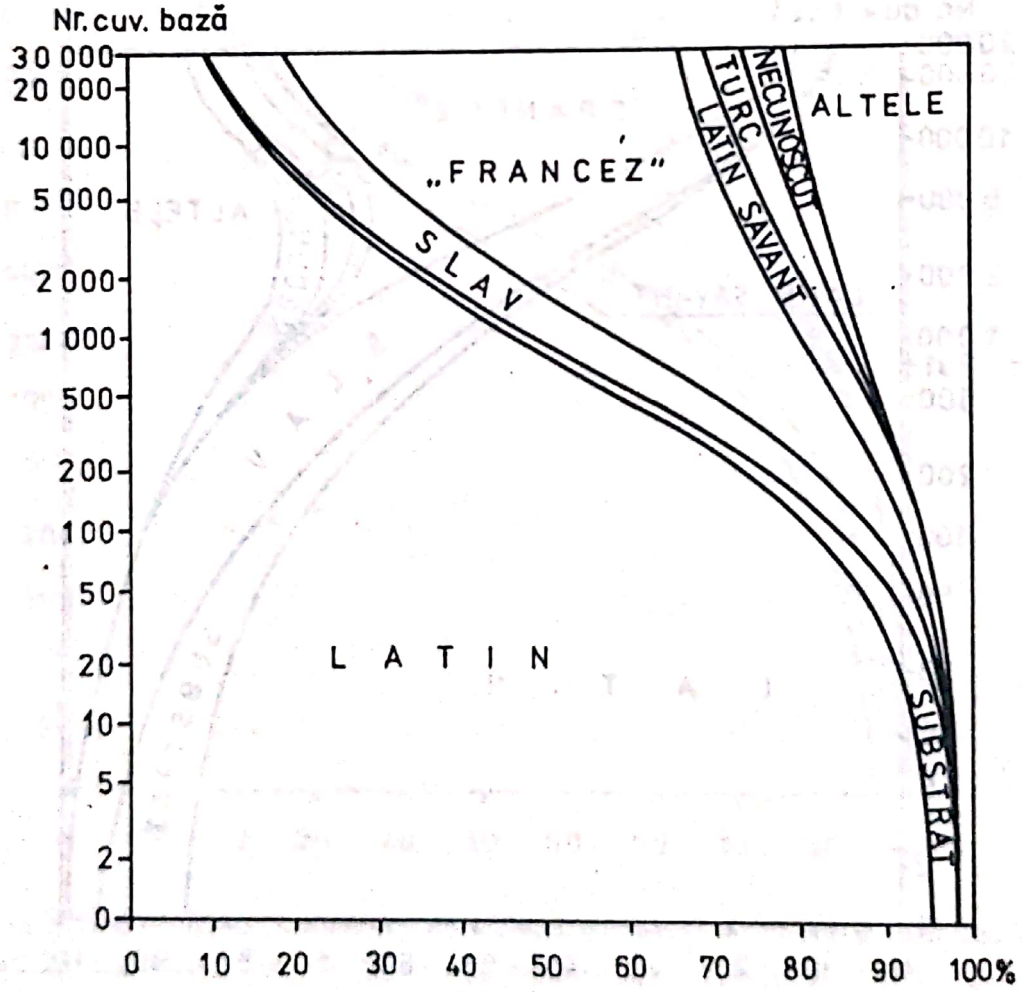
FIG. 20.2 TOPOGRAMA PRODUCTIVITĂŢII



(modelul $M_s$ — fig. 20.3). Ni se relevă şi aici poziţia intermediară a modelului $M_a$, care acordă elementului latin un loc mai însemnat decît cel pe care acesta îl ocupă în $M_s$, dar mai puţin important decît cel deţinut în $M_p$. În virtutea dependenţei de vechime a producţiei de derivate, topograma productivităţii privilegiază substratul şi influenţa slavă veche în raport cu înrîuririle mai recente.

Ar mai fi de notat, pentru preîntîmpinarea nedumeririi eventuale a cititorului, că, avînd în vedere caracterul non-discret al ecuaţiilor utilizate, partea inferioară a graficelor prezintă ciudăţenia consemnării unui număr fracţionar de cuvinte, ceea ce, fireşte, nu are o semnificaţie lingvistică, ci doar o explicaţie de ordin matematic. Potrivit graficului 20.1, cele mai importante 5 cuvinte româneşti provin din latineşte în proporţie de 93,82%, primele 2 în proporţie de 95,11% şi nici chiar cel mai însemnat cuvînt nu e cu totul latin, ci doar într-un procent

FIG. 20.3 TOPOGRAMA POLISEMIEI



de 95,74% (din calcule rezultă că el aparţine în măsură de 3,17% substratului, 0,75% — elementului vechi slav, 0,17% celui bulgaro-sîrbocroat, fiind, în plus, într-o proporţie neînsemnată de provenienţă incertă (0,07%), bulgară (0,04%), maghiară (0,04%) şi onomatopeică (0,01%). Simpla rotunjire a valorilor înlătură paradoxul, procentele reprezentate grafic începînd să aibă sens numai din momentul în care ajung să reprezinte un număr supraunitar de cuvinte.

## 21. Uzajul şi problemele sale

Am explicat deja pe larg (cap. 3) de ce nu ne raliem opţiunii acelor lingvişti pentru care frecvenţa de folosire constituie criteriul hotărîtor de apreciere a importanţei cuvintelor. Argumentele noastre apără, în fond, autonomia structurii interne a limbii de tendinţele de privilegiere a aspectelor pragmatice care, umbrind unitatea sistemului lexical, insistă asupra varietăţii realizărilor lui concrete, ceea ce conduce, în ultimă instanţă, la o pulverizare a entităţii limbii într-un fascicul de ipostaze particulare infinit diversificabile, mergînd de la aşa numitele „stiluri" (beletristic, tehnico-ştiinţific, publicistic, administrativ etc.) pînă la sumedenia de idiolecte ale vorbitorilor individuali. Cu toate acestea, ideea întîietăţii argumentului statistic rămîne la noi adînc înrădăcinată cu atît mai mult cu cît, la prestigiul inerent oricărui criteriu „obiectiv", s-a adăugat şi aureola conferită de serviciile pe care teoria hasdeiană a „circulaţiei cuvintelor" le-a adus luptei pentru afirmarea specificului romanic al lexicului românesc. Acest merit istoric a contribuit în mod esenţial la consolidarea poziţiei metodei inventarelor de frecvenţă în practica lexicologică, pînă în imediata actualitate. E suficient să menţionăm doar faptul, simptomatic, că pînă şi într-o lucrare[1] ai cărei autori exprimă serioase rezerve privind importanţa uzajului pentru stabilirea vocabularului reprezentativ al limbii[2], selecţia celor mai importante unităţi lexicale se întemeiază în cele din urmă preponderent tot pe rezultatele statisticii lingvistice.

---

[1] *Vocabularul reprezentativ al limbilor romanice* (sub red. M. Sala) Ed. Ştiinţifică, 1988.

[2] „... fără să tăgăduim meritele statisticii lingvistice, o considerăm insuficientă pentru a putea conduce la atingerea scopului urmărit de noi..." (op. cit. pag. 13).



Scanned with OKEN Scanner

Astfel, dacă pe criteriile puterii de derivare şi al numărului de sensuri cercetarea citată a reţinut 791 şi respectiv, 1039 de cuvinte româneşti, dicţionarul de frecvenţe a servit la selectarea a nu mai puţin de 1907 cuvinte[1]. Detaşarea de prejudecata priorităţii uzajului, deşi proclamată, nu izbuteşte aşadar să devină efectivă, rămînînd în mare parte un deziderat neonorat.

În ceea ce ne priveşte, am acordat, după cum s-a văzut, întîietate absolută criteriilor *productivitate* şi *polisemie,* neglijînd pînă în momentul de faţă frecvenţa de întrebuinţare a cuvintelor în vorbire şi scriere. Investigaţia noastră ar fi însă incompletă dacă am ignora cu totul acest aspect, cu atît mai mult cu cît esenţa limbii nu poate fi despărţită de finalitatea ei socială care este comunicarea. Regimul vocabularului în procesul de comunicare rămîne în fond o problemă de lingvistică, chiar dacă pe acest teren implicarea (sau, mai exact spus, intricarea) aspectelor extralingvistice este mult mai puternică decît în celelalte sectoare ale studiului limbii.

În raport cu aspectele investigate anterior, pentru care modelarea matematică a reprezentat o veritabilă premieră, uzajul s-a bucurat încă de multă vreme de o tratare în termeni matematici. Sînt peste şapte decenii de cînd stenograful J. B. Estoup a formulat, într-o primă versiune, legea statistică confirmată de G. K. Zipf (1949) şi ameliorată de B. Mandelbrot privind raporturile ce subzistă între rangurile ocupate de cuvinte într-o listă de frecvenţe şi valorile frecvenţelor înseşi. Aceasta şi alte legi statistice stabilite pentru diferite limbi se cer verificate şi în cazul limbii române, căreia nu putem avea certitudinea că i se aplică automat, cum ne-o dovedeşte, de pildă, contraexemplul următor:

Una dintre conjecturile statisticii lingvistice afirmă că, la nivel de dicţionar, între numărul de sensuri $s$ ale cuvintelor şi numărul de cuvinte $n_s$ care posedă $s$ sensuri subzistă relaţia:[2]

$$n_s s^2 = \text{constant} \qquad (21.1)$$

---

[1] Rămîne neexplicată de autori neconcordanţa dintre acest număr, determinat a posteriori, şi opţiunea programatică iniţială conform căreia „criteriul uzajului a fost folosit prin selectarea primelor 2000 de cuvinte din dicţionarul de frecvenţe al lui A. Juilland (*op. cit.* pag. 19). Nedumerirea persistă şi după consultarea listelor întocmite pentru celelalte limbi romanice, în legătură cu care se declară de asemenea că au fost preluate de la Juilland cîte 2000 de cuvinte, care însă în bilanţul final se dovedesc a fi numai 1859 pentru italiană, 1881 pentru franceză şi 1879 pentru spaniolă.

[2] S. Marcus, Ed. Nicolau, S. Stati: *Introducere în lingvistica matematică* Ed. Ştiinţifică, 1966, pag. 231.

După ce am ordonat (v. cap. 16) toate cuvintele din DLRM potrivit numărului lor de înţelesuri, ne este uşor să verificăm valabilitatea acestei „legi". Tabelul 21.1 prezintă în cele 3 linii ale sale valorile parametrilor $s$, $n_s$ şi a funcţiei $n_s s^2$.

*Tabelul 21.1*

| $s$ | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_s$ | 2 | 6 | 6 | 9 | 15 | 16 | 28 | 49 | 76 | 134 | 310 | 623 | 1585 |
| $n_s s^2$ | 450 | 1176 | 1014 | 1296 | 1815 | 1600 | 2268 | 3136 | 3724 | 4824 | 7750 | 9968 | 14265 |

Comentariile sînt de prisos. Nimic mai puţin variabil decît presupusa constanţă a membrului stîng al egalităţii 21.1., dovadă că rezultatele stabilite în condiţiile altor idiomuri nu trebuie extinse asupra limbii române fără o atentă verificare prealabilă.

Tot în cartea citată (pag. 226), este prezentată şi aşa-numita lege „număr-frecvenţă" ce exprimă legătura dintre numărul $n_f$ al cuvintelor de frecvenţă egală dintr-un text dat şi frecvenţa $f$ a acestor cuvinte. Formal, legea seamănă cu precedenta, avînd următoarea expresie:

$$n_f f^2 = C_L$$

unde constanta $C_L$ depinde de lungimea $L$ a textului considerat.

Pentru a testa valabilitatea acestei legi, vom utiliza datele statistice cuprinse în lucrarea pe care Luiza Seche a consacrat-o lexicului artistic eminescian [1]. Ţinînd seamă şi de numele proprii şi de cuvintele străine, inventariate separat (v. tabelul 13 *op. cit.*), care trebuie adăugate listei celor 5016 termeni comuni identificaţi de autoare în ansamblul celor $L = 67\,735$ de cuvinte-text, se obţine situaţia din tabelul 21.2.

*Tabelul 21.2*

| $f$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_f$ | 2156 | 823 | 463 | 287 | 219 | 148 | 123 | 84 | 82 | 57 | 43 | 51 | 47 |
| $n_f f^2$ | 2156 | 3292 | 4167 | 4592 | 5475 | 5328 | 6027 | 5376 | 6642 | 5700 | 5203 | 7344 | 7943 |

| $f$ | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_f$ | 37 | 31 | 26 | 25 | 25 | 18 | 10 | 15 | 19 | 17 | 21 | 11 | 16 |
| $n_f f^2$ | 7252 | 6975 | 6656 | 7225 | 8100 | 6498 | 4000 | 6615 | 9196 | 8993 | 12096 | 6875 | 10816 |

Cum lesne se observă, şi de această dată o conjectură prea uşor acceptată drept lege statistică se vede dezminţită de realitate. Deşi

[1] *Lexicul artistic eminescian în lumină statistică* Ed. Academiei, 1974.

neesențială din punctul de vedere al cercetării noastre, o atare infirmare e de natură să atragă atenția asupra pericolului virtual ca și în cazul altor legi „generale" extrapolarea pripită asupra limbii române să se dovedească incorectă.

În mod special ne interesează aici măsura în care este respectată legea Estoup-Zipf, deoarece, așa cum vom încerca să arătăm mai departe, implicațiile lingvistice ale acestei relații cantitative sînt deosebit de bogate și semnificative.

Începem prin a reaminti enunțul legii în forma ei clasică, fără ajustările ulterioare aduse de B. Mandelbrot, modificări care au sporit indiscutabil precizia modelării fenomenului, dar au și complicat forma sa de exprimare matematică, făcînd extrem de dificile, dacă nu chiar imposibile acele prelucrări și inferențe pe care ne-am propus să le desfășurăm în continuare.

Dacă totalitatea unităților lexicale ce compun un text de lungime $L$ este prezentată sub forma unei liste de cuvinte dispuse în ordinea descrescătoare a numărului lor de ocurențe și dacă se notează cu r rangul unui cuvînt din lista respectivă și cu $f(r)$ frecvența absolută a acestuia, legea Estoup-Zipf afirmă că produsul $rf(r)$ este constant oricare ar fi $r$. Vom avea deci:

$$rf(r) = k$$

sau

$$f(r) = \frac{k}{r} \tag{21.2}$$

Fie $L$ lungimea textului și $N$ numărul total de cuvinte distincte ce îl alcătuiesc. Prin definiție, suma frecvențelor absolute ale tuturor celor $N$ cuvinte este egală cu $L$:

$$\sum_{r=1}^{N} f(r) = k \sum_{r=1}^{N} \frac{1}{r} = L \tag{21.3}$$

Expresia $\sum_{r=1}^{N} \frac{1}{r}$ cuprinde primii $N$ termeni ai seriei armonice a cărei sumă, pentru $N \geq 50$, poate fi aproximată cu formula:

$$\sum_{r=1}^{N} \frac{1}{r} = c + lnN \tag{21.4}$$

unde $c$ este constanta lui Euler, egală cu 0,577215.



Scanned with OKEN Scanner

Introducînd expresia 21.4 în 21.3 se determină valoarea constan- tei $k$:

$$k = \frac{L}{c + lnN}$$

după care relaţia 21.2 capătă forma:

$$f(r) = \frac{L}{r\,(c + lnN)} \tag{21.5}$$

Sîntem acum în măsură să deducem un corolar foarte important al legii Estoup-Zipf. El se referă la problema, de deosebit interes lingvistic, a determinării cotei părţi dintr-un text de lungime $L$ pe care o acoperă cele mai frecvente $n$ cuvinte. Această fracţiune $l$ din $L$ are expresia:

$$l = \sum_{r=1}^{n} f(r) = \frac{L}{c + lnN} \sum_{r=1}^{n} \frac{1}{r}$$

Or, potrivit formulei 21.3:

$$\sum_{r=1}^{n} \frac{1}{r} = c + ln\, n$$

de unde rezultă imediat că:

$$\frac{l}{L} = \frac{c + ln\, n}{c + ln\, N} \tag{21.6}$$

Expresia 21.6 ne permite să calculăm cît loc ocupă dintr-un text dat cele mai frecvente $n$ cuvinte din totalul celor $N$ unităţi care îl compun. Astfel, presupunînd că numărul cuvintelor distincte întîlnite într-un text se cifrează la 1000, putem determina cu ajutorul formulei 21.6 întinderea totală ocupată de primele 10, 50, 100 sau 500 de cuvinte din lista de frecvenţe absolute. În primul caz, de exemplu, vom avea $n = 10$, $N = 1000$ şi:

$$\frac{l}{L} = \frac{c + ln\, 10}{c + ln\, 1000} = \frac{0{,}577 + 2{,}303}{0{,}577 + 6{,}908} = \frac{2{,}880}{7{,}485} = 0{,}3848$$

ceea ce înseamnă că cele mai frecvente 10 cuvinte acoperă nu mai puţin de 38,48% din întregul text. O corecţie este însă necesară întrucît, după cum am mai menţionat, formula 21.5 din care am dedus-o şi pe 21.6 furnizează rezultate suficient de exacte numai pentru valori $n \geq 50$. De fapt, în cazul $n = 10$ putem să nu mai apelăm la



Scanned with OKEN Scanner

soluţia aproximativă oferită de relaţiile de tip 21.4 şi 21.5, deoarece pentru valori mici ale lui $n$ suma primilor $n$ termeni se calculează direct fără dificultate:

$$\sum_{r=1}^{10} \frac{1}{r} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} = 2{,}929$$

de unde rezultă imediat: $\frac{l}{L} = \frac{2{,}929}{7{,}485} = 0{,}3913$

Mai exact spus, cele 10 cuvinte din capul listei de frecvenţe ocupă nu 38,48%, ci 39,13% din totalul textului. Se vede că diferenţa nu e prea importantă. În plus, ea scade rapid odată cu creşterea rangului. Pentru $n = 20$, formula 21.4 furnizează $\sum_{r=1}^{20} \frac{1}{r} = 3{,}573$, pe cînd însumarea directă a celor 20 de termeni ai sumei ne dă $\sum_{r=1}^{20} \frac{1}{r} = 3{,}598..$ Ca procentaj din text, acestea reprezintă 47,74% şi, respectiv, 48,07%, abaterea fiind de două ori mai mică decît în primul caz. Oricum, e de reţinut faptul remarcabil că, practic, jumătate din lungimea unui text ce cuprinde 1000 de unităţi lexicale distincte constă în repetarea a numai 20 de cuvinte foarte frecvente. Concordă acest rezultat, întrucîtva paradoxal, cu datele de observaţie de care dispunem?

E binecunoscut faptul că, în materie de statistică a vocabularului, cercetarea cea mai amplă de care se dispune pentru limba română este cea întreprinsă de un colectiv condus de profesorul Alphonse Juilland de la Universitatea din Stanford (California). În cadrul unui program mai larg de investigare a lexicului limbilor romanice „majore", acesta a elaborat, pe baza unor principii unitare (chiar dacă nu inatacabile), dicţionare de frecvenţă a cuvintelor din limbile spaniolă (1964), română (1965), franceză (1970) şi italiană (nepublicat). Întemeiat pe excerptarea unor texte literare şi ştiinţifice totalizînd 500.000 de ocurenţe lexicale, texte aparţinînd, în principal, perioadei interbelice (deşi, prin excepţie, au fost luate în considerare şi piesele lui Caragiale sau, dincolo de cealaltă extremitate a intervalului, romanul *Desculţ* de Zaharia Stancu), *Dicţionarul de frecvenţe al cuvintelor româneşti*[1] prezintă nu puţine slăbiciuni şi scăpări. În mod

[1] A. Juilland, P.M.H. Edwards, Ileana Juilland: *Frequency Dictionary of Rumanian Words*, Mouton & Comp, London, The Hague, Paris. 1965.



Scanned with OKEN Scanner

inexplicabil, sînt omise cu totul cuvinte de largă întrebuinţare (prepoziţia *cu*, articolul *cel, cea*, numeralul *unu, una*), în timp, ce alţi termeni apar dublu înregistraţi (adverbul *astfel* figurează atît la rangul 125, cu 382 de ocurenţe, cît şi pe poziţia 343 ÷ 6 cu 143 de ocurenţe). Clasamentul întocmit este viciat şi de considerarea drept cuvinte unice a unor locuţiuni alcătuite din elemente de sine stătătoare: *pentrucă* (?), *dupăce*(?!), *deaceia*(??!). Comentatorii români ai lucrării s-au arătat, în general, necruţători faţă de aceste carenţe, declarînd dicţionarul inutilizabil, ceea ce nu ni se pare adevărat decît în parte. Încercînd să fim drepţi, vom spune că această vastă lucrare statistică rămîne *statistic* valabilă, în sensul că deficienţele sale privesc mai ales elemente şi aspecte individuale şi că *în mare* ea ne spune totuşi cîteva lucruri semnificative şi, pînă la un punct, corecte despre regimul statistic al vocabularului limbii române. Bunăoară, în problema care ne interesează aici, scăpările de detaliu ale autorilor dicţionarului nu sînt de natură să afecteze grav situaţia globală. În ciuda absenţei fortuite a cîtorva cuvinte importante, cota parte din text reprezentată de ocurenţele celor mai frecvente 1000 sau 2000 de unităţi lexicale nu e de crezut să fi suferit alterări grave, aşa încît utilizarea datelor de acest fel furnizate de dicţionar ne pare justificată. A. Juilland precizează, de pildă, în comentariul ce premerge listelor de frecvenţă propriu-zise, că numărul însumat de ocurenţe ale cuvintelor ce ocupă primele 500 de ranguri este de 323.141, valoare care pentru următoarele grupe de cîte 500 de ranguri scade, respectiv, la 35.405, 20.592, 13.729, 10.020, 7.638, 5.915, 4.632, 3.612 şi 2.323 de apariţii. De asemenea, menţionează că cele mai frecvente 10, 50 sau 100 de cuvinte totalizează, respectiv, 119.997, 213.667 şi 244.609 ocurenţe. Din păcate, verificarea pe care am efectuat-o a arătat că, fie la adunare, fie la transcrierea listei de frecvenţe trebuie să se fi strecurat unele erori, căci numărul însumat de ocurenţe determinat de noi nu a coincis cu cel declarat de Juilland şi, în plus, ecartul înregistrat nu a fost acelaşi în toate cazurile, ceea ce nu se poate explica decît prin existenţa mai multor greşeli în diverse zone ale bilanţului. Astfel, adunînd frecvenţele din dicţionar, am găsit pentru primele 50 de cuvinte 216.939 de ocurenţe în loc de 213.667 (o diferenţă, deci, de 3.272), pentru primele 100—248.420 de ocurenţe în loc de 244.609 (o diferenţă de 3.811), iar pentru primele 500—326.946 de ocurenţe în loc de 323.141 (o diferenţă de 3.805). Neconcordanţele sînt, fireşte, supărătoare, dar nu într-atît de grave încît să compromită cu totul tentativa noastră de valorificare selectivă a datelor furnizate de dicţionar. Diferenţele sesizate reprezintă un procent relativ mic din totalul cuvintelor-text, ponderea lor scăzînd cu cît ne referim la ranguri

mai mari, adică exact acolo unde aplicarea corolarului menţionat al legii lui Zipf este cea mai îndreptăţită. Pentru testarea valabilităţii acestui corolar, am sistematizat în tabelul 21.3 datele furnizate de dicţionarul lui Juilland astfel: pentru diferite ranguri $n$, cuprinse în limitele $n = 1$ — $n = 5.000$, adică pentru întregul spectru lexical nominalizat în dicţionar, sînt prezentate frecvenţele absolute însumate ale cuvintelor pînă la rangul respectiv (linia a treia), precum şi cota procentuală reprezentată de aceste frecvenţe din totalul celor 500.000 de cuvinte-text considerate de Juilland (linia a patra).

*Tabelul 21.3*

| $n$ | 1 | 2 | 5 | 10 | 20 | 30 | 50 | 100 | 200 |
|---|---|---|---|---|---|---|---|---|---|
| $ln\ n$ | 0 | 0,69 | 1,61 | 2,30 | 3,00 | 3,40 | 3,91 | 4,61 | 5,30 |
| $f$ | 19748 | 36433 | 73744 | 119997 | 167643 | 191527 | 216939 | 248420 | 281220 |
| % | 3,95 | 7,29 | 14,75 | 24,00 | 33,53 | 38,31 | 43,39 | 49,68 | 56,24 |

| $n$ | 500 | 1000 | 1500 | 2000 | 2500 | 3000 | 4000 | 5000 |
|---|---|---|---|---|---|---|---|---|
| $ln\ n$ | 6,21 | 6,91 | 7,31 | 7,60 | 7,82 | 8,00 | 8,29 | 8,52 |
| $f$ | 326946 | 362351 | 382943 | 396672 | 406692 | 414330 | 424897 | 430832 |
| % | 65,39 | 72,49 | 76,59 | 79,35 | 81,34 | 82,87 | 84,98 | 86,17 |

Reprezentate grafic în figura 21.1, aceste date generează o curbă a cărei alură va trebui comparată cu cea a funcţiei 21.6 particularizată pentru $N = 20.000$ (din păcate, nu cunoaştem exact cîte cuvinte distincte a găsit A. Juilland în materialul excerptat, dar, după propriile sale declaraţii, ar fi vorba de vreo 20.000, poate ceva mai multe, întrucît se pare că s-au eliminat numele străine[1]):

$$\frac{l}{L} = \frac{0{,}577 + ln\ n}{10{,}48}$$

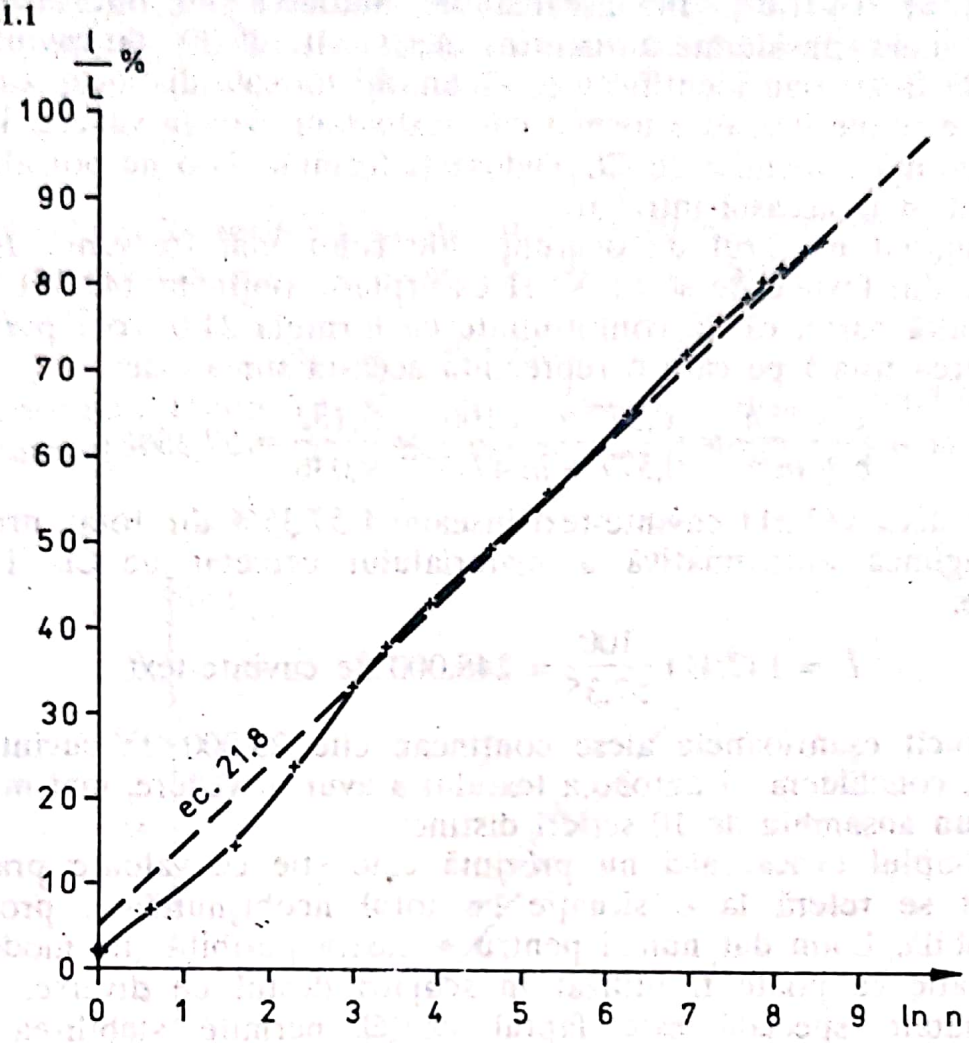
Funcţia $\frac{l}{L}(n)$ este, după cum se vede, liniară. Dreapta respectivă (trasată punctat în fig. 21.1) se suprapune aproape perfect începînd

[1] Reprezentarea logaritmică atenuează simţitor imprecizia decurgînd din această necunoaştere, dat fiind că, la valori mari ale argumentului, logaritmul creşte foarte încet: faţă de $ln\ 20.000 = 9{,}90$ avem, de pildă, $ln\ 21.000 = 9{,}95$, ceea ce înseamnă că, la scara desenului, o diferenţă de 1000 de cuvinte-titlu nu reprezintă decît vreo jumătate de milimetru.



Scanned with OKEN Scanner

FIG. 21.1




cu $ln\, n = 3$ (ceea ce înseamnă $n = 20$) peste curba „experimentală" furnizată de dicţionar. Ştiut fiind faptul că, oricum, formula stabilită de noi (21.6) nu e valabilă întrutotul decît pentru $n \geq 50$ (adică $ln\, n = 3{,}91$, punct unde coincidenţa teoretic-experimental este perfectă), putem considera legea 21.6 drept validată de realitatea lingvistică a limbii române. Odată stabilit acest lucru, avem, în sfîrşit, dreptul să mergem mai departe, urmărind consecinţele ce decurg din confirmarea relaţiei 21.6.

Dacă, de obicei, lungimea totală $L$ a textului este cunoscută, pot surveni uneori situaţii în care, dintr-o pricină sau alta, să nu dispunem de date asupra acestui parametru. Iată un exemplu:

Întreprinzînd o cercetare statistică a lexicului limbii române vechi, în intenţia de a determina vocabularul de bază al acesteia, Claudia Tudose[1] a excerptat, între altele, şi un număr de texte redactate în

[1] *Vocabularul fundamental al limbii române vechi* în „Sistemele limbii" Ed. Academiei, Bucureşti, 1970, pag. 119—164.

secolul al XVII-lea. În lucrarea sa, autoarea ne informează că eşantioanele considerate au cuprins fiecare cîte 25.000 de cuvinte-text şi că, în final, s-au identificat 4.717 unităţi lexicale distincte. Ceea ce nu ni se spune însă este tocmai cîte texte s-au avut în vedere. Listele de frecvenţă întocmite de Cl. Tudose şi formula 21.6 ne pot ajuta să răspundem la această întrebare.

Însumînd numărul de ocurenţe ale celor mai frecvente 100 de cuvinte din textele de secol XVII excerptate, obţinem 142.311. Ştim, pe de altă parte, că, în conformitate cu formula 21.6, cota parte din întinderea totală pe care o reprezintă această sumă este:

$$\frac{c + \ln n}{c + \ln N} = \frac{0{,}577 + \ln 100}{0{,}577 + \ln 4717} = \frac{5{,}182}{9{,}036} = 57{,}35\%.$$

Dar dacă 142.311 cuvinte-text înseamnă 57,35% din total, urmează că lungimea aproximativă a materialului cercetat de Cl. Tudose este de:

$$L = 142.311 \frac{100}{57{,}35} = 248.000 \text{ de cuvinte-text}$$

Întrucît eşantioanele alese conţineau cîte 25.000 de cuvinte-text fiecare, conchidem că autoarea textului a avut în vedere, mai mult ca sigur, un ansamblu de 10 scrieri distincte.

Exemplul evocat aici nu prezintă cine ştie ce valoare practică, întrucît se referă la o situaţie cu totul neobişnuită şi, probabil, irepetabilă. L-am dat numai pentru a ilustra posibilităţile modelului matematic ce poate fi utilizat în scopuri destul de diverse. De o însemnătate specială este faptul că el permite stabilirea unei corespondenţe-standard între lungimea unui text şi numărul cuvintelor distincte care îl alcătuiesc. Abaterile de la această corelaţie sînt importante fiindcă se pretează la interpretări de ordin stilistic. În ce constă problema?

Relaţia 21.5, pe care o transcriem aici pentru a facilita urmărirea raţionamentului:

$$f(r) = \frac{L}{r(c + \ln N)} \qquad (21.5)$$

exprima, pentru un text de lungime L dată, legătura dintre rangul unui cuvînt şi frecvenţa acestuia. Particularizînd $r = N$, vom putea scrie:

$$f(N) = \frac{L}{N(c + \ln N)}$$

relaţie care ne furnizează frecvenţa absolută a celui mai puţin folosit termen din textul în cauză. Or, practica generală a despuierii textelor ne arată că există întotdeauna cel puţin un cuvînt (de fapt nu unul,



Scanned with OKEN Scanner

ci o întreagă categorie, cum vom vedea, chiar cea mai numeroasă dintre toate[1]) care nu are decît o singură ocurenţă. Al $N$-lea cuvînt din inventarul statistic are deci frecvenţa $f(N) = 1$. Introducînd $r = N$ şi $f(N) = 1$ în 21.5, obţinem:

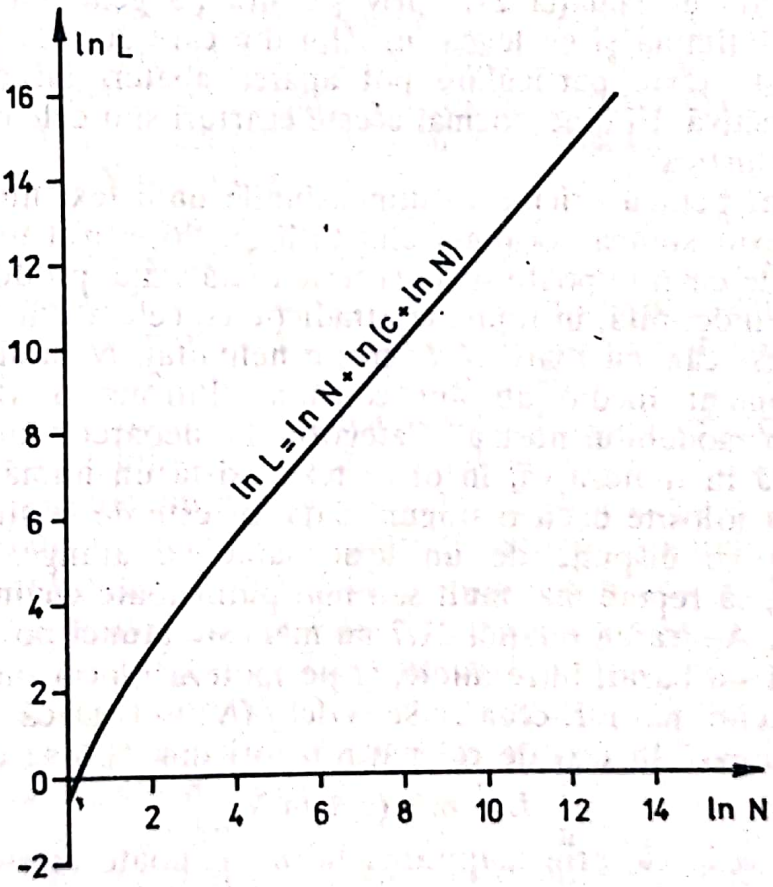
$$L = N(c + \ln N) \tag{21.7}$$

Formula 21.7 exprimă legătura dintre lungimea unui text şi numărul cuvintelor distincte ce-l compun. Trebuie notat că forma ei logaritmată:

$$\ln L = \ln N + \ln(c + \ln N) \tag{21.8}$$

nu constituie o relaţie liniară, cum s-a crezut[2]. Totuşi, datorită faptului că termenul al doilea al sumei din membrul drept creşte mult mai

FIG. 21.2




[1] De exemplu, dintre cele 5.016 cuvinte extrase de L. Seche din textele eminesciene, un număr de 2.014 nu apar decît o singură dată.

[2] În lucrarea de pionierat *Introducere în lingvistica matematică* de S. Marcus, Ed. Nicolau şi S. Stati (Ed. Ştiinţifică, 1966) se prezintă o expresie empirică liniară de forma $\ln L = \frac{\ln N - \ln a}{\alpha}$ cu $a$, $\alpha$ = constant, verificată pe texte de Shakespeare, Lessing, Puşkin, Rimbaud.

încet decît primul, curba de ecuaţie 21.8 prezintă o tendinţă marcată de liniarizare (v. fig. 21.2), ceea ce face ca, practic, începînd cam cu $N = 50$ ea să poată fi considerată o dreaptă. Dat fiind că, de fapt, relaţiile stabilite de noi prezintă, după cum s-a mai arătat, o precizie mulţumitoare de-abia pentru valori $N \geq 50$, se poate afirma că nu există nici o contradicţie între rezultatul nostru şi formula citată în nota de subsol, cele două funcţii fiind numeric compatibile, în ciuda structurii lor analitice diferite.

Odată stabilită relaţia 21.7 care exprimă legătura nemijlocită dintre întinderea unui text şi numărul cuvintelor utilizate *în medie* la alcătuirea lui, sîntem în posesia unui instrument eficace de evaluare a bogăţiei mijloacelor lexicale folosite de un autor. Am subliniat complementul pentru a evidenţia faptul că interdependenţa dintre $L$ şi $N$ reflectată de ecuaţia 21.7 priveşte situaţia generală a limbii (la fel cum se întîmpla şi cu legea lui Zipf din care aceasta a derivat) şi că în diferite texte particulare pot apărea abateri însemnate de la regula respectivă. Ei bine, tocmai aceste ecarturi sînt cele care prezintă relevanţă stilistică.

E evident pentru oricine că dimensiunile unui text nu dictează cu necesitate extensiunea vocabularului utilizat. Posesorul unui bagaj de numai 500 de cuvinte poate vorbi şi scrie toată viaţa, producînd „texte" de lungime indefinită, în totală contradicţie cu cele afirmate de relaţia 21.7. În acest caz, cu toate că $L$ creşte nelimitat, $N$ rămîne de la un anumit moment încolo absolut constant. Infirmă o atare situaţie valabilitatea modelului nostru? Categoric nu, deoarece formula 21.7 a fost stabilită în ipoteza că, în orice text, există un număr de cuvinte care nu sînt folosite decît o singură dată, or este de la sine înţeles că o persoană ce dispune de un lexic sărac va ajunge, pălăvrăgind incontinent, să repete mai mult sau mai puţin toate cuvintele pe care le cunoaşte. Aplicarea relaţiei 21.7 nu mai este atunci posibilă, fiindcă stabilirea ei s-a bazat, între altele, şi pe ipoteza unicei întrebuinţări a cuvîntului celui mai rar, ceea ce se scrie: $f(N) = 1$. Dacă orice unitate lexicală figurează în text de cel puţin $m$ ori din 21.6 se deduce:

$$L = mN(c + \ln N)$$

de unde se vede că, prin majorarea lui $m$, $L$ poate creşte şi el oricît de mult sau, altminteri spus, că repetînd aceleaşi cuvinte se poate vorbi sau scrie oricît, fără ca prin aceasta să fie periclitată valabilitatea legii Estoup-Zipf.

Mai mult însă decît libertatea de a flecări avînd acoperirea legilor lingvisticii statistice, ne interesează aici problema bogăţiei vocabularului unor scriitori despre care s-au făcut, nu o dată, afirmaţii grăbite, bazate pe impresii superficiale sau intuiţii neverificate. Astfel,



Scanned with OKEN Scanner

un autor cu privire la care s-a spus că posedă un lexic relativ sărac este însuşi Eminescu. Să vedem cum stau cu adevărat lucrurile.

Analizînd antumele eminesciene şi 10 dintre postume, Luiza Seche a găsit în cuprinsul a 34.703 cuvinte-text un număr de 3.670 de cuvinte-titlu. Rezolvarea prin încercări a ecuaţiei 21.7 pentru $L = 34.703$ conduce la $N = 3.921$, valoare superioară celei constatate de autoarea citată. Totalul cuvintelor folosite de Eminescu în versurile cercetate este deci inferior numărului expectat, ceea ce ar pleda pentru o oarecare modicitate a lexicului poetului. E drept însă că diferenţa nu e prea mare, valoarea reală reprezentînd cca. 93,6% din cea scontată. Asemănătoare este şi situaţia prozei artistice eminesciene, unde de-a lungul a 33.748 de cuvinte-text aceeaşi cercetătoare a întîlnit 3.530 de cuvinte distincte, adică 92,4% din cele 3.822 anticipate de formula 21.7. Spre comparaţie, să menţionăm şi rezultatul obţinut de Mircea Goga în inventarierea poeziei lui Octavian Goga[1]: 3.188 de cuvinte-titlu (din care 3.054 termeni comuni şi 134 de nume proprii) pe parcursul a 39.971 de cuvinte-text. În acest din urmă caz, ar fi fost de aşteptat, potrivit ecuaţiei 21.7, să se găsească 4.431 de unităţi lexicale distincte. Numărul celor efectiv prezente se mărgineşte, aşadar, la numai 71,95% din valoarea expectată, contrastul cu situaţia vocabularului eminescian fiind izbitor. Dacă nu ne putem pronunţa încă (vom vedea mai încolo de ce) asupra sărăciei sau bogăţiei lexicului lui Eminescu, avem motive temeinice să afirmăm că, în comparaţie cu el, Goga foloseşte un vocabular mult mai sărac.

Să continuăm însă examinarea zestrei lexicale eminesciene. Tot Luiza Seche a stabilit că numărul total de cuvinte-titlu utilizate în proza şi în poeziile investigate este de 5.234. Lungimea însumată a textelor considerate reprezintă, potrivit autoarei studiului, 68.451 de cuvinte-text. Introducînd această ultimă valoare în relaţia 21.7, obţinem rezultatul surprinzător că, per total, în textele eminesciene ar fi trebuit să figureze nu doar 5.234 ci 7.233 de cuvinte distincte. Iată că acum vocabularul scriitorului, ce părea de proporţii mijlocii, se vădeşte dintr-odată aproape la fel de modest ca acela al lui O. Goga: el s-a redus la numai 72,4% din valoarea scontată. Ce să înţelegem de aici?

Încercînd să „salvăm" prestigiul de mînuitor de limbă al Luceafărului poeziei româneşti, am putea pune fenomenul observat pe seama unităţii stilistice a scrisului eminescian. Prozatorul Eminescu nefiind decît o altă ipostază a poetului, era firesc ca între vocabularul prozei şi cel al poeziei sale să fie mult mai multe elemente comune decît în cazul producţiilor a doi autori diferiţi, ceea ce poate explica

---

[1] *Lexicul si structura stilului în poezia lui Octavian Goga* Ed. Minerva, 1989.



Scanned with OKEN Scanner

satisfăcător discrepanţa dintre densitatea cvasinormală de cuvinte-titlu din cele două secţiuni ale operei sale considerate separat şi imaginea, cu mult deosebită, pe care ne-o oferă abordarea globală a creaţiei literare a scriitorului.

Să nu ne grăbim însă cu recurgerea la atari sofisme întru „apărarea" poetului, înainte de a privi lucrurile şi din alt unghi. Explorări statistice ale scrisului românesc s-au mai întreprins, cum bine se ştie, şi pentru colecţii eteroclite de texte, purtînd pecetea unor personalităţi de facturi adesea foarte diferite. Mai mult decît atît, corpusurile cercetate proveneau de multe ori din zone disjuncte ale culturii şi ştiinţei, din presă, drept, administraţie, Or, dacă şi pentru asemenea ansambluri de scrieri, aflate sub semnul eterogenului şi al diversităţii, se va constata aceeaşi tendinţă de folosire a unui număr de cuvinte mai mic decît cel standard, prezis de modelul matematic, atunci va trebui să admitem că ceea ce s-a luat drept o expresie a sărăciei lexicului eminescian nu este decît reflexul unei situaţii cu caracter mai general, a cărei explicaţie e de căutat în altă parte.

În ceea ce priveşte literatura beletristică, cel mai reprezentativ studiu statistic întemeiat pe un material cu profil antologic este cel întreprins de membrii Cercului de lingvistică matematică al Secţiei de fonetică şi dialectologie din cadrul Institutului de lingvistică din Bucureşti, ale cărui rezultate au fost publicate în 1959 sub semnătura lui V. Şuteu în revista *Studii şi cercetări lingvistice,* X, nr. 3 la paginile 419—445. S-au avut atunci în vedere texte aparţinînd unui număr de 10 poeţi, prozatori şi dramaturgi reprezentativi din prima jumătate a veacului nostru. Pe întinderea celor 50.000 de cuvinte-text ce alcătuiau corpusul considerat, au fost identificate 4.547 de cuvinte-titlu distincte. Cifra prognozată de formula 21.7 este însă sensibil mai mare, căci pentru $L = 50.000$ rezultă $N = 5.447$. Numărul real reprezintă aşadar numai 83,48% din cel scontat, deşi de data aceasta amprenta stilistică a unui singur autor lipseşte. S-ar putea vorbi cel mult de o unitate de alt nivel, proprie limbii literaturii române din acea etapă istorică în ansamblul ei.

Materialul utilizat de A. Juilland la elaborarea dicţionarului său se întinde însă şi în afara acestor limite. Pe lîngă opere beletristice, el mai include articole de ziar, lucrări de medicină, botanică, inginerie, agronomie etc. Şi totuşi... Printre cele 500.000 de cuvinte-text s-au identificat numai vreo 20.000 de termeni distincţi, adică mai puţin de jumătate din cele 44.000 de unităţi la care aveam dreptul să ne aşteptăm în virtutea relaţiei 21.7! Se poate desprinde o idee generală din aceste observaţii în fond disparate?



Scanned with OKEN Scanner

Pentru a încerca să ne apropiem de un răspuns, am reprezentat pe diagrama 21.3, alături de graficul funcţiei $ln\,N = f(ln\,L)$ cele trei puncte corespunzătoare datelor oferite de statisticile menţionate: punctul $S$ (10,82; 8,42), ce reflectă relaţia dintre lungimea şi bogăţia lexicală a textelor studiate în lucrarea lui V. Şuteu, punctul $E$(11,13; 8,62) furnizat de cercetarea Luizei Seche asupra lexicului eminescian şi punctul $J$ (13,12; 9,90) rezultat din statistica lui A. Juilland. Se observă că aceste trei puncte nu sînt dispuse oricum, ci ele par să se situeze pe o aceeaşi curbă, tangentă la „simili-dreapta" ce ilustrează grafic funcţia $ln\,N = f(ln\,L)$. Faptul e de natură să sugereze că relaţia cvasiliniară dintre ln N şi ln L e valabilă numai într-un domeniu limitat atît la stînga (ceea ce ştiam), cît şi la dreapta (ceea ce aflăm acum). Existenţa limitării la dreapta nu trebuie să surprindă. Ea este perfect explicabilă dacă ne gîndim că orizontul lexical al oricărui producător de texte ca şi cel al oricărui grup de asemenea producători rămîne finit, în vreme ce capacitatea acestora de a genera texte nu cunoaşte, practic, frontiere. Este de aceea de aşteptat ca, începînd de la un anumit punct încolo, panta curbei $ln\,N = f(ln\,L)$ să scadă treptat, funcţia prezentînd aspectul unei curbe logistice, adică tinzînd asimptotic către o orizontală ce marchează plafonul corespunzător orizontului lexical respectiv.

Fără pretenţia de a oferi soluţii definitive, încercăm să sugerăm un mod posibil de abordare a problemei. O curbă care să îndeplinească simultan condiţia de a trece prin punctele $S$, $E$ şi $J$ şi de a admite, în plus, o asimptotă orizontală e cea descrisă de ecuaţia:

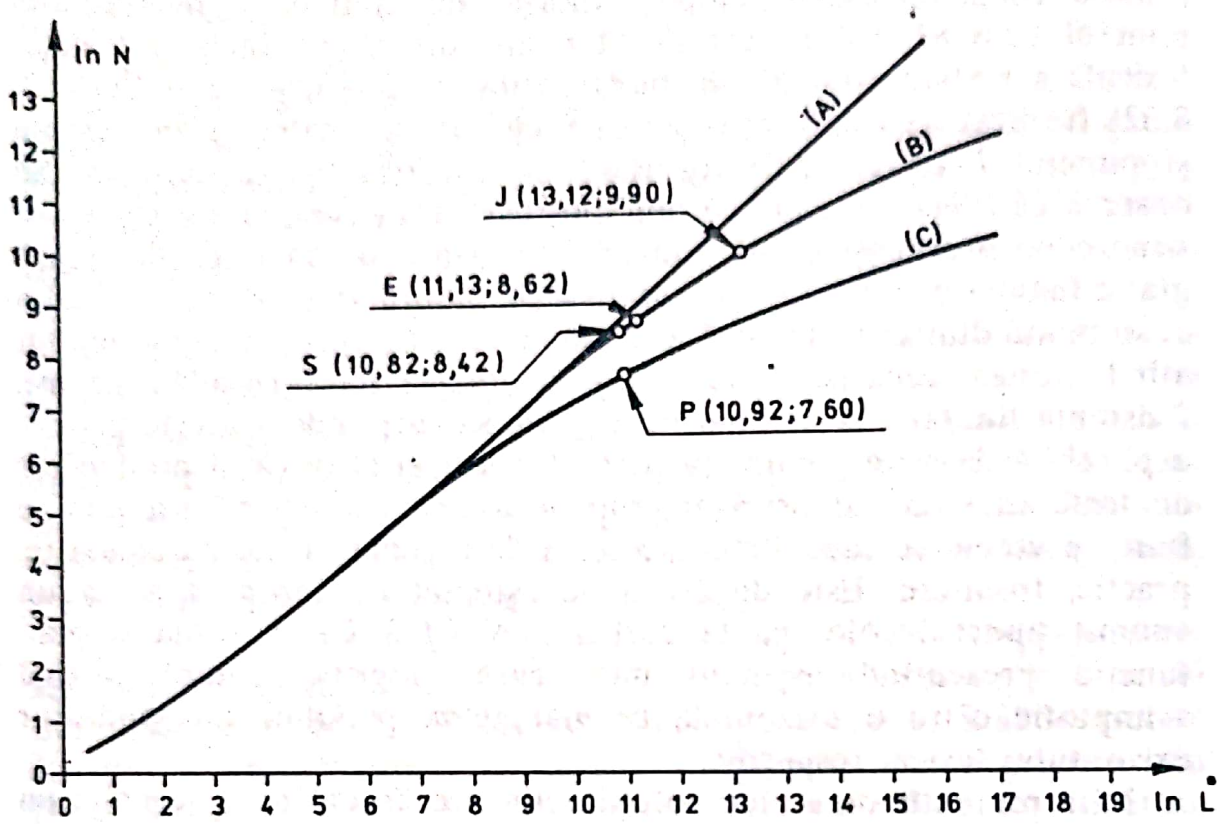
$$ln\,N = 15{,}62 - 21{,}2\ exp\ (-0{,}1\ ln\,L) \qquad (21.9)$$

Reprezentată în figura 21.3 (curba B), ea este aproape tangentă (la scara desenului, dacă nu şi în litera definiţiei matematice, ceea ce ar fi fost imposibil, căci presupunea impunerea prea multor condiţii unei funcţii cu numai 3 parametri) la cvasidreapta descrisă de ecuaţia 21.8. Funcţia 21.9 se desprinde aşadar de aceasta din urmă, îndulcindu-şi urcuşul, fapt ce reflectă creşterea progresivă a decalajului dintre numărul cuvintelor-titlu la care ne-am fi aşteptat în virtutea relaţiei 21.7 şi al celor indicate, mai realist, de ecuaţia 21.9 pentru texte de lungimi $L$ echivalente. Astfel, dacă într-un text conţinînd 100.000 de ocurenţe lexicale numărul de cuvinte distincte rezultate din expresiile 21.8 şi 21.9 este, respectiv, de 10.200 şi 7.400, într-unul de 1 milion de ocurenţe e de cca 94.000 şi, respectiv, 31.000 de cuvinte, iar într-unul de 10 milioane de ocurenţe — de cca. 710.000 şi, respectiv, 88.500 de cuvinte.



Scanned with OKEN Scanner

FIG. 21.3




Nu e greu de observat că al doilea şir de valori e cu mult mai plauzibil decît cel dintîi. Cu cît cercul se extinde, cu atît devine mai evident că realitatea lingvistică se conformează mai degrabă unei legi de tipul 21.9 decît uneia de forma 21.8. Ca o situaţie extremă, să presupunem, de exemplu, cazul unei biblioteci de 10 milioane de volume, cuprinzînd în medie cîte 100.000 de cuvinte fiecare. Lungimea totală a „textului" reprezentat de acest uriaş (dar nu imposibil de întîlnit, căci există atari biblioteci) tezaur de carte va fi de $L = 10^{12}$ cuvinte-text. Calculele arată că, potrivit relaţiei cvasiliniare 21.8 deduse din legea Estoup-Zipf, în cărţile respective va trebui să găsim 40 de miliarde (!) de cuvinte distincte, ceea ce e, în mod evident, o absurditate, întrucît atîtea cuvinte nu există nici măcar în toate limbile pămîntului luate la un loc (fiecare dintre cele vreo 4000 de limbi cunoscute ar trebui să dispună de cîte 10 milioane de cuvinte!) În schimb, presupunînd că în loc de 21.8 se respectă formula 21.9, nu mai e vorba decît de numărul absolut acceptabil, de 1 milion şi jumătate de cuvinte.



Scanned with OKEN Scanner

Ceea ce nu trebuie, în nici un caz, uitat este faptul că estimările de mai sus privesc orizontul lexical al limbii române literare din prima jumătate a secolului nostru. Înscrierea lexicului artistic eminescian pe o aceeaşi curbă cu inventarele lexicale cu profil atît de divers întocmite de V. Şuteu şi A. Juilland demonstrează că limba poetului, pe nedrept suspectată de o relativă paupertate, este, dimpotrivă, foarte bogată. Varietatea ei suportă comparaţia cu universul lexical al ansamblului literaturii, nu numai beletristice, scrise într-o perioadă cînd vocabularul limbii române se primenea spectaculos prin infuzia (ca să nu spunem invazia) masivă a elementului neologic.

Pentru a înţelege mai bine proporţiile acestui proces şi, deopotrivă, amplitudinea reală a lexicului eminescian, care anticipează trăsăturile şi tendinţele întregului vocabular al limbii culte contemporane, vom recurge la comparaţia cu o situaţie din trecutul mai îndepărtat, respectiv de la începuturile limbii române scrise.

Printre cercetările de statistică a vocabularului desfăşurate la noi în ultimele trei decenii, se numără şi studiul Vioricăi Pamfil consacrat unui monument literar al secolului al XVI-lea: *Palia de la Orăştie*[1]. Reţinem din lucrarea publicată de această autoare numai cuplul de valori care ne interesează: $L = 55.197$, $N = 2.003$. Aşadar numai 2000 de cuvinte diferite într-un corpus mai vast decît cel considerat de V. Şuteu, în care acesta numărase, totuşi, peste 4500 de unităţi lexicale distincte. După cum era de aşteptat, punctul $P$ (10,92; 7,60), corespunzător vocabularului *Paliei*, de pe diagrama din figura 21.3 e situat cu mult sub celelalte, semn că avem într-adevăr de a face cu un lexic foarte sărac.

Procedînd ca mai sus, putem căuta expresia unei exponenţiale cu asimptotă orizontală care să fie deopotrivă tangentă la curba $ln\, N = f(ln\, L)$ dată de 21.8 şi să treacă prin punctul $P$. Ecuaţia rezultantă este:

$$ln\, N = 12{,}60 - 14{,}85\, exp(-0{,}1\, ln\, L) \qquad (21.10)$$

iar curba ce o ilustrează grafic e cea notată cu (C) în figura 21.3.

În intenţia de a face paralela mai sugestivă, vom calcula cu ajutorul formulei 21.10 cîte cuvinte distincte ar trebui să conţină, în perspectiva orizontului lexical al *Palliei,* un text cu lungimea de o sută de mii, un milion, zece milioane şi, respectiv o mie de miliarde (echivalentul

---

[1] *Contribuţii la studiul statistic al lexicului românesc. Structura lexicală a Paliei de la Orăstie* Studii de lexicologie, vol. I, Universitatea „Babeş-Bolyai", Cluj, 1965, pag. 157—232.



Scanned with OKEN Scanner

bibliotecii cu zece milioane de volume) de cuvinte. Rezultatele sînt: 2650, 7000, 15000 şi, respectiv, 118000 de cuvinte-titlu, în loc de 7400, 31000, 88500 şi, respectiv, 15000000 cîte rezultaseră pentru limba secolului XX. Comentariile sînt inutile. Se cuvine însă menţionat că a judeca bogăţia şi varietatea limbii române din veacul al XVI-lea după profilul lexical al *Paliei,* traducere destul de căznită a unui text „specializat", ar constitui o mare nedreptate. Limba vie, care de-abia începînd cu răspîndirea cărţilor coresiene începea să graviteze pe nesimţite către o expresie literară unificatoare, trebuie să fi fost de o luxuriantă varietate, despre care puţinele documente scrise ce s-au păstrat nu ne pot da nici cea mai palidă idee.

Revenind acum la o problemă pe care am mai amintit-o în treacăt la începutul capitolului de faţă şi anume cea a legăturii dintre frecvenţa absolută şi totalitatea cuvintelor ce înregistrează numărul respectiv de ocurenţe într-un text dat, vom încerca să stabilim o expresie matematică a acestui raport, sprijinindu-ne pe cele demonstrate pînă în prezent. În vederea acestui scop, reluăm relaţia 21.5:

$$f(r) = \frac{L}{r\,(c + \ln N)} \tag{21.5}$$

Vom nota cu $n_1$, $n_2$, $n_3$, ..., $n_f$, ... numărul de cuvinte care au o singură şi, respectiv, 2, 3, ..., $f$, ... apariţii în textul considerat. În lista frecvenţelor, numărul de ordine al ultimului element este, firește, $N$. După cum s-a arătat, numărul $N$ poate fi calculat ca funcţie de $L$ impunînd în expresia 21.5 condiţia $f(N) = 1$:

$$f(N) = \frac{L}{N\,(c + \ln N)} = 1$$

de unde

$$\frac{L}{c + \ln N} = N \tag{21.11}$$

În mod asemănător, rangul $r\,(2)$ al ultimului cuvînt care figurează în text exact de două ori se obţine din condiţia $f(r) = 2$:

$$2 = \frac{L}{r\,(2)\,(c + \ln N)}$$

sau, ţinînd seamă de 21.11:

$$r(2) = \frac{N}{2}.$$



Scanned with OKEN Scanner

Diferența dintre $N$ și $r\,(2)$ va fi exact numărul $n\,(1)$ al cuvintelor cu o singură ocurență:

$$n_1 = r\,(1) - r\,(2) = N - \frac{N}{2} = \frac{N}{2}$$

În continuare, $r\,(3)$ se deduce din:

$$3 = \frac{L}{r\,(3)\,(c + ln\,N)}, \quad r\,(3) = \frac{L}{3(c + ln\,N)} = \frac{N}{3}$$

și devine posibilă determinarea lui $n_2$:

$$n_2 = r\,(2) - r\,(3) = \frac{N}{2} - \frac{N}{3} = \frac{N}{6}$$

și așa mai departe.

În cazul general:

$$n_f = r\,(f) - r\,(f + 1) = \frac{N}{f} - \frac{N}{f+1} = \frac{N}{f\,(f+1)},$$

de unde:

$$n_f f\,(f + 1) = N$$

sau

$$n_f f^2 + n_f f = \text{constant} \qquad (21.12)$$

Relația 21.12 constituie forma, în sfîrșit, corectă a legii număr-frecvență, a cărei exprimare am arătat (v. supra pag. 314) că fusese încercată cu ajutorul ecuației:

$$n_f f^2 = \text{constant} \qquad (21.13)$$

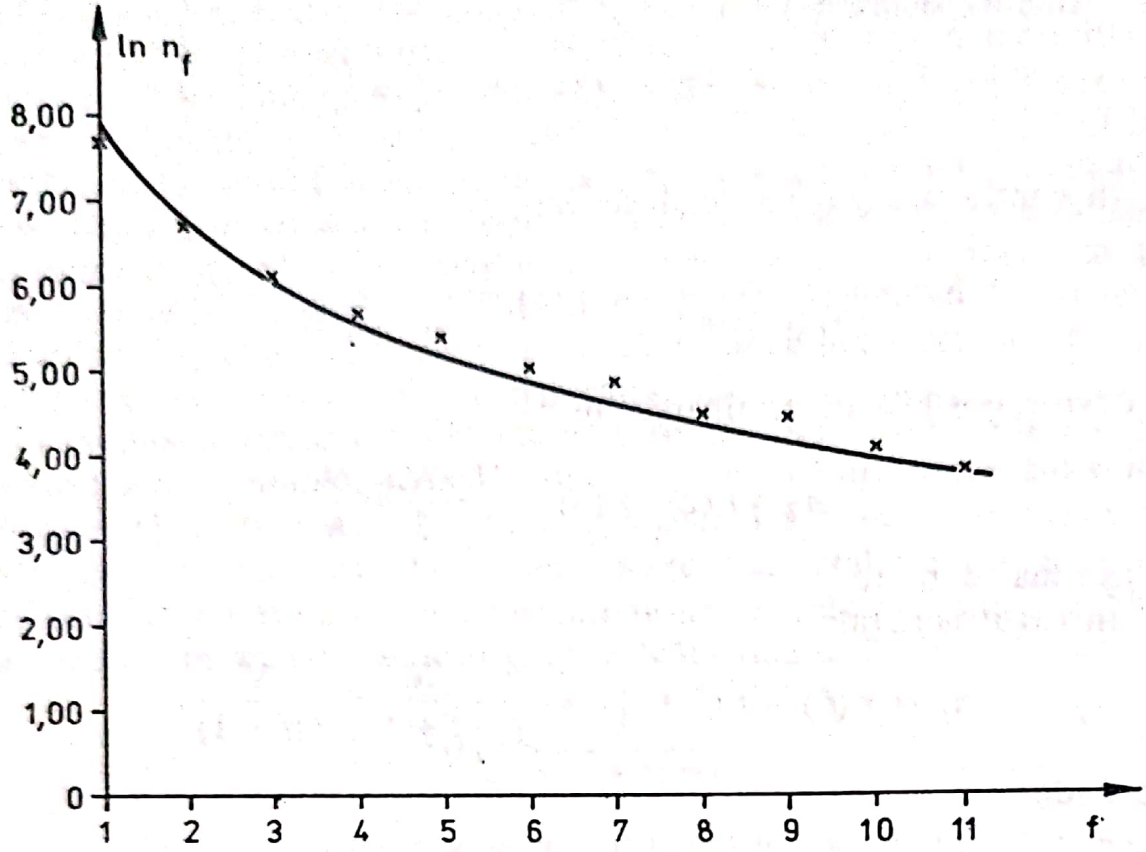
Înțelegem acum de ce rezultatele statisticii eminesciene efectuate de Luiza Seche nu verificau relația 21.13; din membrul drept al egalității lipsea termenul $n_f f$. În același timp, ne explicăm și cum s-a putut strecura eroarea: formula 21.13 a fost testată probabil pentru frecvențe mari, adică acolo unde valoarea $f$ poate fi neglijată în raport cu $f^2$, ceea ce nu mai e acceptabil în zona frecvențelor reduse.

Că expresia 21.12 modelează convenabil raporturile număr-frecvență ne-o arată graficul din figura 21.4, unde s-a reprezentat în coordonate semilogaritmice (pentru a diminua amplitudinea variației lui $n_f$, a cărui valoare ajunge să depășească 2000, pentru $f = 1$) relația dintre $n_f$ și $f$ în corpusul cercetat de Luiza Seche. Se vede că punctele



Scanned with OKEN Scanner

FIG. 21.4




experimentale urmăresc îndeaproape graficul trasat pe baza relaţiei 21.12, dovadă că legea dedusă de noi se mulează satisfăcător pe realitatea lingvistică.

O ultimă observaţie privitoare la relaţia 21.7 se referă la implicaţiile acesteia într-un domeniu oarecum îndepărtat de obiectul studiului de faţă şi anume învăţarea limbilor străine. Pentru cine deprinde o limbă citind şi traducînd cu ajutorul dicţionarului, prezintă interes ritmul de apariţie a cuvintelor noi. E de presupus că, odată ajuns la a $n$-a pagină a unei cărţi, cititorul nu va mai fi nevoit să facă apel la dicţionar la fel de des ca la începutul lecturii. Pe măsură ce avansează, tot mai multe cuvinte i-au devenit cunoscute şi, ca urmare, frecvenţa consultării dicţionarului scade. Cum putem descrie în termeni cantitativi acest proces?

Să presupunem de exemplu, că numărul mediu al cuvintelor de pe fiecare pagină a cărţii este de 400. În ipoteza că bagajul lexical iniţial de care dispune cititorul e nul, se pune întrebarea de cîte ori va trebui el să recurgă la dicţionar pentru a traduce corect prima pagină. Răspunsul ni-l dă relaţia 21.7, care pentru $L = 400$ furnizează $N = 80$, ceea ce înseamnă că pe parcursul primei pagini vor fi întîlnite cca. 80



Scanned with OKEN Scanner

de cuvinte distincte. Trecînd la pagina următoare, este de aşteptat ca numărul termenilor necunoscuţi să scadă. În ce proporţie?

Într-un text de două pagini, cuprinzînd, conform ipotezei noastre, 800 de ocurenţe lexicale, aceeaşi formulă 21.7 ne arată că trebuie să existe $N = 144$ de cuvinte-titlu. Cum 80 figurau în prima pagină, rezultă că cea de a doua nu mai conţine decît 64. Numărul apelurilor la dicţionar s-a redus cu 16. Într-un mod similar (calculînd $N$ pentru $L = 1\,200$) aflăm şi de cîte ori trebuie recurs la dicţionar pentru a citi pagina a treia a cărţii. Rezultatul este $N = 60$. Scăderea este, conform aşteptărilor, monotonă, dar nu şi uniformă. Procesul suferă o încetinire treptată, altminteri am ajunge, împotriva intuiţiei şi a logicii, după numai cîteva pagini, la cunoaşterea perfectă a limbii cu pricina. Astfel, încheind primele 100 de pagini, pe parcursul cărora am făcut cunoştinţă cu 44 456 de cuvinte, mai întîlnim la pagina 101 încă 38 de cuvinte necunoscute. Acest rezultat pare destul de puţin încurajator pentru învăţăceii în ale limbilor străine, dar lucrurile s-ar putea să nu stea tocmai aşa. Vocabularul nici unui autor nu coincide cu întregul lexic al limbii şi, ca atare, nici ritmul de apariţie a cuvintelor noi într-o carte nu trebuie gîndit în funcţie de situaţia din limbă în general. Exemplul *Paliei de la Orăştie* demonstrează că un autor se poate descurca, la rigoare, şi cu un număr mult mai mic de unităţi lexicale, în care caz frecvenţa noilor apariţii se vede sensibil diminuată. Problema se pretează la dezvoltări interesante, care, nu îşi au însă locul aici. E momentul să închidem o şi aşa prea lungă paranteză pentru a reveni la problema centrală a cărţii de faţă, care este caracterizarea profilului etimologic al vocabularului limbii române prin intermediul cîtorva parametri lingvistici cuantificabili dintre care, în acest capitol, ne concentrăm atenţia asupra frecvenţei de folosire a cuvintelor.

Din păcate, pentru respectivul indicator nu avem la dispoziţie un fond de date suficient de complet privind situaţia din limba română contemporană ca să putem purcede la prelucrări de genul celor efectuate în cazul parametrilor productivitate şi polisemie. Am menţionat deja faptul că autorii celor mai de seamă lucrări româneşti de statistică lexicală (Bolocan, Maneca, Şuteu), fie din economie de spaţiu editorial, fie din pricină că urmăreau obiective diferite de al nostru, nu au încredinţat tiparului lista integrală a cuvintelor întîlnite în texte şi a frecvenţelor de apariţie a acestora. Nici A. Juilland nu o face, deşi cele 5000 de unităţi pe care le consemnează ar fi, în principiu, suficiente pentru tragerea unor concluzii, dar lacunele, inconsecvenţele, chiar identificarea eronată a unor cuvinte (bunăoară:



Scanned with OKEN Scanner

absenţa prepoziţiei *întru* din rîndul cuvintelor româneşti frecvent utilizate indică o regretabilă confuzie cu paronimul *între* în contexte ca *într-un, într-o, într-adevăr* etc.) viciază raporturile dintre componentele etimologice ale vocabularului, afectînd grav credibilitatea rezultatelor.

Am preferat de aceea să ne adresăm din nou statisticii datorate Luizei Seche care, pe de o parte, ne oferă atît de necesarul tablou complet al spectrului de frecvenţe din corpusul considerat, iar pe de alta, suscită un plus de interes întrucît se referă la o paradigmă stilistică emblematică pentru limba română literară din toate timpurile. Dacă raporturile structurale perene din sînul vocabularului românesc sînt de căutat în opera unui autor concret, atunci nu încape îndoială că acesta nu poate fi decît Eminescu, artistul preocupat deopotrivă de punerea în valoare a farmecului limbii din „cronice bătrîne" ca şi de promovarea necomplexată, creatoare, a neologismului percutant.

Urmînd calea deja cunoscută din capitolele precedente, va trebui să repartizăm pe origini cele 5016 cuvinte distincte clasificate de autoarea studiului în ordinea numărului lor de ocurenţe, sarcină, din nefericire, nu mult uşurată de precizările etimologice din carte, care sînt în destule cazuri greşite[1].

Am reexaminat aşadar etimologiile şi am regrupat materialul lexical pe criteriul provenienţei, menţinînd pentru fiecare origine în parte ordinea relativă dată de numărul de ocurenţe.

În cazul elementului latin moştenit rezultatul acestor operaţii este lista următoare:

| | |
|---|---|
| 1. el (ea) | 3058 |
| 2. şi (conj.) | 2370 |
| 3. de (prep.) | 2334 |
| 4. în | 2017 |

[1] E cel puţin ciudată atribuirea originii latine unor derivate de la rădăcini slave vechi ca *atîrna, desclesta, despletit, goliciune, iubit, lipit, mlădia, sclipitor, trăsnet* sau unor cuvinte de provenienţă bulgărească *(puf)*, neogreacă *(măcar)* ori sîrbocroată *(odor)*. Extravaganţa de a acorda unor termeni derivaţi un statut etimologic diferit de cel al cuvintelor-matcă face curate ravagii: verbul *vicleni* e socotit slav, deşi *viclean* este diagnosticat corect drept maghiar, *boltit* poartă menţiunea „latin" cu toate că provine de la *boltă*, explicat prin sîrbocroată, *blîndeţe* e considerat slav, în timp ce lui *blînd* i se recunoaşte originea latină. Alte atribuiri sînt de-a dreptul năucitoare: despre *ţărînă* aflăm că ar proveni din franţuzeşte, ca şi *sol* (= mesager), confundat poate, din viteză, cu nota muzicală omonimă. Şi, fiindcă veni vorba de notă, derivatul *notiţă* poartă menţiunea „etimologie necunoscută"!



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 5. | un (o) | 1516 |
| 6. | fi | 1450 |
| 7. | cu | 1251 |
| 8. | avea | 1111 |
| 9. | şie (se) | 1092 |
| 10. | pe | 1022 |
| 11. | să | 957 |
| 12. | eu | 954 |
| 13. | tu | 906 |
| 14. | al (a) | 823 |
| 15. | ce | 806 |
| 16. | nu | 799 |
| 17. | din | 587 |
| 18. | la | 531 |
| 19. | ca | 529 |
| 20. | vrea | 497 |
| 21. | că | 469 |
| 22. | care | 407 |
| 23. | mai (adv.) | 392 |
| 24. | cînd | 364 |
| 25. | meu | 307 |
| 26. | vedea | 303 |
| 27. | ochi | 299 |
| 28. | cel (cea) | 273 |
| 29. | unu (una) | 270 |
| 30. | tău | 248 |
| 31. | prin | 235 |
| 32. | lume | 233 |
| 33. | putea | 209 |
| 34. | cum (conj.) | 206 |
| 35. | zice | 192 |
| 36. | viaţă | 184 |
| 37. | umbră | 183 |
| 38. | tot (pron.) | 182 |
| 39—40. | a, între | 181 |
| 41. | tot (adj.) | 179 |
| 42. | faţă | 176 |
| 43. | dulce | 173 |
| 44. | părea | 171 |
| 45. | trece | 169 |
| 46. | veni | 168 |



Scanned with OKEN Scanner

47. lună . . . . . 163
48. mînă . . . . . 161
49. şti . . . . . 159
50—51. face, noapte . . . . . 158
52. şi (adv.) . . . . . 157
53. ca . . . . . 154
54—55. alb, numai . . . . . 151
56. noi . . . . . 141
57. acest . . . . . 138
58—60. mare, negru, suflet . . . . . 134
61. fără . . . . . 132
62—63. pentru, vis . . . . . 128
64. da (vb.) . . . . . 126
65. inimă . . . . . 125
66. cer . . . . . 121
67. cap . . . . . 117
68—69. acel, frumos . . . . . 116
70. stea . . . . . 115
71. floare . . . . . 114
72—73. sta, zi . . . . . 112
74—75. lung, său . . . . . 111
76. încet . . . . . 110
77. om . . . . . 108
78. atunci . . . . . 106
79. lumină . . . . . 105
80. lăsa . . . . . 104
81. pînă . . . . . 103
82—83. frunte, peste . . . . . 97
84—85. dacă, pămînt . . . . . 95
86—87. păr, uita . . . . . 93
88. spune . . . . . 92
89—90. apă, mare (subst.) . . . . . 91
91—92. atît, duce . . . . . 90
93—94. gură, voi (pron.) . . . . . 88
95—96. apoi, pune . . . . . 87
97—99. braţ, doi, după . . . . . 86
100. nimic . . . . . 84
101—104. adînc, cădea, moarte, sub . . . . . 82
105—106. auzi, muri . . . . . 79
107—108. acum, tot (adv.) . . . . . 78
109—111. nor, plin, simţi . . . . . 77



Scanned with OKEN Scanner

112—113. minte, rază . . . . . 76
114—115. rămîne, unde . . . . . 75
116—119. aur, cal, cine, vînt . . . . . 74
120—121. codru, lui (art.) . . . . . 73
122—123. fereastră, lîngă . . . . . 72
124—126. fată, înger, întreg . . . . . 71
127. bate . . . . . 70
128—129. albastru, ridica . . . . . 69
130—131. soare, unul . . . . . 68
132. spre . . . . . 67
133. lacrimă . . . . . 66
134—137. arunca, carte, rece, vechi . . . . . 65
138—139. loc, singur . . . . . 64
140. înalt . . . . . 62
141. lua . . . . . 61
142—143. nici, roşu . . . . . 60
144—146. casă, nostru, plînge . . . . . 59
147—149. ajunge, cela, mult . . . . . 58
150. mic . . . . . 57
151—156. an, aşa, cale, ci, întinde, ţinea . . . . . 56
157—161. aer, cuprinde, încă, sîn, trist . . . . . 55
162. dor . . . . . 54
163—168. alt, altul, azi, femeie, piatră, piept . . . . . 52
169—170. între, mie . . . . . 51
171—172. mult (adj), tînăr . . . . . 50
173—178. acesta, asupra, ăst, bătrîn, bine, pătrunde . . . . 49
179—180. ăsta, deschide . . . . . 48
181—185. ieşi, intra, întreba, nici, urmă . . . . . 47
186—189. frunză, împărat, suna, sus . . . . . 46
190—194. acela, arăta, lin, nimeni, vreun . . . . . 45
195—199. blînd, foc, pas, uita, umple . . . . . 44
200—204. căuta, marmură, părete, răsări, stinge . . . . . 43
205—215. argint, asculta, crede, creşte, lac, mijloc, mişca, pădure, pierde, purta, rîde . . . . . 42
216—218. merge, săruta, tremura . . . . . 41
219—220. dînsul, începe . . . . . 40
221—228. cînta, cît, dormi, însă, întoarce, rupe, somn, şedea 39
229—230. arde, fugi . . . . . 38
231—236. aşeza, domn, printre, rege, sau, vale . . . . . 37
237—244. asemenea, cum, mamă, munte, picior, scrie, umăr, undă . . . . . 36



Scanned with OKEN Scanner

245—248. aripă, Dumnezeu, înţelege, jos . . . . . 35
249—251. ascunde, chiar, rîu. . . . . 34
252—255. margine, oare, prinde, sară . . . . . 33
256—261. apropia, lucru, ori, tată, ţară, zbura . . . . . 32
262—268. cît, cîte, colo, cuvînt, drept, parte, verde . . . . . 31
269—270. aproape, însumi . . . . . 30
271—278. atîta, dintru, geană, îmbrăca, înălţa, naşte, nume, strînge . . . . . 29
279—288. abia, amîndoi, bun, cere, înainte, înaintea, închide, înconjura, pieri, popor . . . . . 28
289—295. aştepta, bătrîn, fiinţă, întuneric, mormînt, mort, pleca . . . . . 27
296—305. acoperi, aduce, astăzi, curge, deget, farmec, negură, oară, poartă, timp . . . . . 26
306—310. chema, cîntec, culca, ramură, uşor . . . . . 25
311—324. adormi, aduna, afară, cărare, dori, galben, nişte, pace, răspunde, scutura, senin, ureche, viu, vostru 24
325—333. cînt, fir, limpede, luci, masă, nou, petrece, umbla, voi . . . . . 23
334—341. adevăr, bucată, iarbă, pasăre, părinte, ruga, semn, vînăt . . . . . 22
342—351. acolo, bine, cald, doar, fund, limbă, luceafăr, moale, subţire, tei . . . . . 21
352—356. ceară, cunoaşte, sări, tăia, vară . . . . . 20
357—364. albină, atinge, lega, mîne, plăcea, sălbatic, şapte, tare . . . . . 19
365—374. adînc, afla, aici, crud, drept, pleca, soartă, suspina, tăcea, trage . . . . . 18
375—389. barbă, creştet, foaie, greu, întuneca, joc, minune, pricepe, puţin, scăpa, scoate, scula, sunet, umed, uşor . . . . . 17
390—398. alege, cît, genunchi, îmbăta, lege, lumina, scară, scînteia, tînăr . . . . . 16
399—411. apărea, aprinde, foarte, fulger, iarnă, învăţa, larg, lemn, lung, luptă, rău, rugăciune, striga . . . . . 15
412—431. aspru, bea, cerc, cîmp, cunună, des, desface, dezmierda, duios, îngîna, juca, june, linişte, preface, răpi, rău (adj.), sînge, sui, toarce, ţărm 14
432—453. alerga, apuca, aramă, arbor, către, cîmpie, cruce, fecioară, ierta, înceta, îneca, legăna, mîna, mut,



Scanned with OKEN Scanner

păsa, pînă, ploaie, spaimă, spate, toamnă, turbure, veghea . . . . . 13

454—471. asemenea (adj.), cam, cît (pron.), deşert, drac, gheaţă, încerca, răsuna, rău (subst.), rotund, schimba, scînteie, sec, spumă, trei, ţese, vărsa, zeu 12

472—486. adăoga, aluneca, amar, au, cetate, corn, cufunda, cugeta, curînd, drept, înşela, jur, preot, sînt, vin 11

487—503. amar (subst.), arc, călare, cot, desprinde, geme, gratie, împărţi, împinge, începea, lupta, minciună, mîngîia, oaste, revărsa, spînzura, trimite . . . . . 10

504—532. adică, apăsa, apune, astupa, bărbat, biserică, capăt, cerca, drept, fier, fierbinte, fior, greier, înclina, lucra, lumînare, miazănoapte, miez, minţi, orb, rar, repede, rumen, senin, smulge, suferi, suspin, urca, veşted . . . . . 9

533—565. aşterne, blestema, cenuşă, coardă, coastă, creier, cuib, cutremura, dimineaţă, faptă, frig, greu, iapă, încărca, închina, întîi, lature, mătasă, mort, necheza, neted, număra, nuntă, roată, ruşine, scîndură, semăna, sete, spată, strîmt, trunchi, vad, vînă . . . . . 8

566—613. alunga, armă, călca, cel, coadă, creţ, doamnă, ferice, fiu, flacără, frate, freamăt, frumos, greu, ieri, îngheţa, însul, întîmpla, lapte, lung, măr, meu, mînca, nas, nimic, pană, pată, patru, păpuşă, preţ, primbla, rătăci, repede, rouă, scaun, sfîşia, sînt (subst.), soarbe, sprînceană, strat, tare, teme, uni, urla, văl, veşmînt, vinde, vultur . . . . . 7

614—656. adăpa, anina, apăra, apleca, atrage, călcîi, căpăta, cărbune, curte, descoperi, desfăşura, deştepta, fie, fiică, greutate, împlini, încheia, încinge, îndemna, îndura, îngust, învia, lepăda, luntre, mira, muşchi, număr, păcat, piuă, pîrîu, primăvară, rîs, salbă, sat, scălda, sterp, subsuoară, suge, tindă, ungher, urî, usca, viţă . . . . . 6

657—706. asămăna, baie, bătaie, bucium, credinţă, crunt, cui, despica, făptură, făt, feri, fura, ger, gros, în, încăleca, încet, înflori, întîi, lăuda, legătură, lună, luneca, mărgea, măsură, mustaţă, muta, pai, piele, plai, plumb, pradă, pulbere, putred, ros, săgeată,



Scanned with OKEN Scanner

scurge, scut, spăla, strecura, supăra, şterge, şuiera, tău, teamă, tîmplă, tort, treaptă, ucide, venin . . 5

707—773. albastru, albi, alina, ameţi, beşică, brumă, cerb, certa, cireş, cît, coase, copt, cumpăra, cumplit, deşert, dezbrăca, dinte, fapt, fîntînă, fraged, frăţîne, fugă, fumega, furnică, împăca, înalt, înăuntru, îndărăt, îngreuia, înota, jura, lătra, lut, măsură, mînie, muia, muşca, muşchi, oală, os, paşte, păgîn, pereche, plop, pretutindeni, puţin, rod, român, sarcină, săpa, seca, soră, sparge, spin, spinteca, strîmb, stuf, sufla, şoarece, tîrziu, trăda, tufă, ţînţar, uda, vargă, vîna . : . . . . . 4

774—855. abate, amează, arşiţă, aşa, aţă, boare, bun, bute, cari, cămaşă, căpătîi, ceaţă, cîne, coace, corb, crăpa, culme, drege, durea, fag, fecior, fierbe, frămînta, frînge, fum, funie, gemet, genune, grîu, ist, închega, îndărătul, înfăşura, însura, învenina, judecată, lat, laţ, lînă, măiestru, mărunt, miel, miere, mişel, muc, mugi, muscă, muşunoi, nepot, nicăiri, ogradă, păstor, pătură, ploua, pom, purice, ram, rană, rădăcină, răni, rîndunea, român (subst.), săptămînă, scădea, schimb, scurt, serba, stîng, tîrziu, turbura, turmă, ud, unge, urcior, urzi, vas, vecin, viers, vioară, volbură, zăcea, zău . . . 3

856—955. ac, acăţa, adia, afunda, aiurea, ajutor, asuda, asurzi, atîta, broască, caier, căpăţînă, căprior, cărunt, căsca, căţel, coamă, coif, creştin (subst.), creştin (adj.), cuc, culege, cununa, cuptor, curcubeu, cute, cutreiera, cuveni, dos, fagură, fiară, fiere, fîn, flămînd, foale, frîu, fulg, furcă, ghem, grindină, grîne, gusta, iederă, ins, împila, împresura, încuia, îndupleca, înfige, jug, junghia, larg, mesteca, moară, muche, nea, ninge, nod, nuc, oaie, oaspe, ospăta, palmă, pîne, pîntece, pui, punte, rărunchi, roade, rost, salcie, sarbăd, saţ, sălta, sămînţă, sătul, sătura, său (subst.), semăna, socru, soţ, speria, subţia, sudoare, şarpe, şase, şea, tătîn, turba, turna, umfla, unchi, veghe, verde, vie, vindeca, vită, vulpe, zbate, zece . . . . . 2

956—1104. adăpost, ager, alb, alinta, alună, amesteca, amuţa, amuţi, arin, aţîţa, ăl, beat, berbece, boteza, bou,



Scanned with OKEN Scanner

bour, buget, capişte, car, căprioară, cearcăn, cercel, cerceta, cerne, cheie, chiar, cina, cină, cinci, cireaşă, ciumă, cînepă, cîştiga, corn, cornut, creastă, creţ, cuteza, cuţit, defăima, depăna, descheia, desculţa, descînta, dezgheţa, disculţ, dulce, dura, falcă, famen, faur, făină, fereca, fiastru, fin, foame, fragă, fuma, fus, ghindă, gras, grăunte, gust, iepure, inel, împrumuta, încinge, înghiţi, îngreuna, înturna, jos, judeca, judeţ, locustă, lup, mai, măduvă, mărita, măruntaie, mătură, mătuşă, merinde, mierlă, muced, mucos, muiere, nară, nouă, nun, ospăţ, papură, par, păşune, pedestru, peşte, piedică, plăcintă, plăpînd, prim, pumn, purcede, rage, răpune, răşină, rîpă, roş, rug, rugăminte, sac, săcure, sănătos, scălda, scăpăta, scărpina, scorbură, scriptură, scuipa, secera, secetă, solz, spic, spinare, spîn, spulbera, spurca, stoarce, sulcină, surd, şele, şes (subst.), şes (adj.), toartă, toca, treiera, tureatcă, turtă, tuşi, ţîţă, unghi, unghie, urgie, uşor, vecin, vierme, viespe, vitreg, vînător, vîntura, vîrtos . . . . . 1

Lexicul eminescian investigat cuprinde aşadar 1 104 cuvinte-bază de origine latină, ceea ce reprezintă 22,01% din totalul unităţilor lexicale distincte (e vorba, evident numai de numele comune) identificate în corpusul considerat. Pentru comparaţie, reamintim că la nivelul dicţionarelor de limbă cu o cuprindere de ordinul a 50 000 de termeni (categorie din care fac parte Dicţionarul „Cartea Românească" al lui I. A. Candrea, DLRM şi DEX) proporţia reprezentată de cuvintele nederivate moştenite din latineşte e de numai 3—4%. Şi mai semnificativ pentru poziţia în limbă a fondului latin primordial al vocabularului românesc în general şi eminescian în special este faptul că „în circulaţie", adică din punctul de vedere al ocurenţelor lexicale (cuvinte-text), clasa cuvintelor-bază de provenienţă latină acoperă 76,89% din total (respectiv 52 124 de ocurenţe din 67 790). La acestea vin să se adauge derivatele formate în româneşte de la termenii-matcă moşteniţi. Iată-i, în ordinea frecvenţei lor absolute:

1. durere . . . . . 71
2. visa . . . . . 48
3. şir . . . . . 34



Scanned with OKEN Scanner

4. întunecos . . . . . 30
5—6. putere, sărutare . . . . . 29
7. simţire . . . . . 28
8. cugetare . . . . . 26
9—11. străbate, strălucit, urma . . . . . 24
12—14. cîntare, închis, liniştit . . . . . 23
15—19. dată, dorinţă, fericire, întins, trecut . . . . . 22
20. deschis . . . . . 21
21—22. străluci, uscat . . . . . 20
23—24. durere, frumuseţe . . . . . 19
25—28. întunecat, miros, nebun, scrisoare, suflare . . . . 18
29—30. fericit, plăcere . . . . . 17
31—35. fire, înţeles, luminos, numi, stare . . . . . 16
36—37. tinereţe, urmări . . . . . 15
38—39. singurătate, tăcere . . . . . 14
40—44. împărătească, luciu, mărire, somnoros, ură . . . . 13
45—52. adînci, adîncime, măreţ, mînuţă, mulţime, nemărginire, repezi, viitor . . . . . 12
53—57. adevărat, depărtare, pierdut, stins, trecător . . . 11
58—67. aurit, fermecat, lungi, nemărginit, nespus, netezi, omenesc, puternic, tărie, trecut . . . . . 10
68—81. argintos, avere, curat, împărăţie, însemna, însenina, înşira, mişcător, ninsoare, răcoare, singuratic, speriat, tremurător, uitare . . . . . 9
82—92. apus, argintiu, ars, doilea, îndoi, înfiora, înflorit, înfunda, înţelept, omenire, visător . . . . . 8
93—120. apăsat, arătare, bărbie, blestem, căldură, ceresc, dreptate, dulceaţă, guriţă, întîmplare, întrebare, înţelepciune, lumesc, luminat, mişcare, muritor, părere, plăcut, plîns, recădea, strălucire, străveziu, sufletesc, tremurînd, turbare, unit, zbor, zburător 7
121—157. amorţit, aprins, avînta, cuvios, despărţi, domnie, frunziş, intrare, înălţime, încetişor, îndoială, îndrepta, înnebuni, întors, învăţat, luciu, mănuşă, mers, mirare, nebunie, neînţeles, nepăsător, neputinţă, oştire, păşi, peliţă, petrecere, plîngere, răguşit, răsfrînge, recunoaşte, spart, suferinţă, tăcut, vechime, vestejit, visare . . . . . 6
158—201. aducere, ales, amărăciune, arcat, beţie, blîndeţe, cădere, colea, dezlega, domni, domnişoară, frăţie, fulgera, împărătesc, început, încreţit, îndărătnic,




Scanned with OKEN Scanner

îngenunchea, îngheţat, înmărmuri, înmiit, însetat, minunat, murind, nebuni, nefericit, nemurire, nemuritor, nevăzut, orbi, plecat, presăra, răsturna, răsufla, reveni, roti, singurel, sperios, spumega, suprafaţă, tristeţă, tunet, ţărînă, vedere . . . . . 5

202—256. afumat, aiura, amestec, arzător, ascuţit, călător, călătorie, căzut, cenuşiu, cerşi, chemare, cuget, cunoscut, firesc, îmbrăţişare, îmbrăţişat, împiedica, încălzi, încreţi, încununa, înduioşare, înecat, înnegri, înstelat, întindere, întristare, învăţătură, învechit, lunatic, măreţie, mări, mărime, molatic, necunoscut, nimici, numire, păta, port, reapărea, retrage, rîzînd, rosti, rugă, rupt, strîns, şeptelea, tras, trăsătură, trăsură, treilea, tremur, undoi, urît, văratic, verdeaţă . . . . . 4

257—347. acoperămînt, adormitor, amestecat, apropiere, binişor, bisericesc, călători, căpeţel, căsătorie, cetăţuie, dătător, despărţire, deştept, dezmierdare, doritor, duioşie, dumnezeiesc, ferici, fulgerător, furiş, îmbătat, înălbi, încetinel, încheietură, încîntat, încrede, încredinţat, îndulci, înfăţişa, îngălbenit, îngeresc, îngheţ, înmuia, însera, înţelegere, învălui, înţelegere, învăli, înveninat, înverzi, jucător, laudă, lăcrimioară, leagăn, lucire, luminiş, mărie, mititel, mîngîiere, mînia, muiat, nedrept, nefiinţă, nemişcat, nepătruns, nins, ochelari, păsărea, plîns, plînsoare, răceală, răsărit, răsfira, rătăcit, rătăcitor, răzbunare, reda, revedea, rîzător, roşeaţă, rugător, ruginit, ruşinos, sărbătoare, scriere, semiîntuneric, simţ, soţie, strălucitor, strîmba, strîngere, supărat, supus, treacăt, trecere, turburat, ucigător, uitat, umbri, umbrit, venire, viitor, vînat . . . . . 3

348—540. acuşi, adîncit, aiurire, albăstriu, alburiu, alintat, amăgit, amuţit, apucat, apunere, arare, ascuns (subst.), ascuns (adj.), asemănare, aţinea, auros, auz, călări, călător, căutare, căutătură, căzînd, cerşitor, cîştig, cununie, cuprins, curgător, cutremur, datori, degeţel, depărtat, descoase, deşteptare, detunat, dezgustat, dizmeţi, dorit, dormind, dumnezeie, făţarnic, femeiesc, femeieşte,



Scanned with OKEN Scanner

fermecător, friguros, fugar, fugit, furişa, gălbui, gheţos, grecesc, greoi, ieşire, îmbătător, îmblînzi, îmbrăcăminte, îmbrăţoşa, împăcat, împietrit, înainta, încăpere, încercare, închegare, închinăciune, încins, încuia, îndoit, îndurare, înflăcărat, înflorire, înfocat, înlătura, înmărmurit, înmiire, înmormîntare, însoţi, înspumat, întineri, întrece, întrecere, întreprinde, întrerupe, întreţese, întristat, întunecime, înviere, învineţi, înzestra, jucărie, junie, junime, limpezi, limpezime, liniştire, lucitor, lucrare, măiestri, măiestrit, mărgini, mesean, micuţ, mincinos, mîncare, mînecă, mucigăit, naştere, necopt, nefăcut, neguros, neîmpăcat, neîndurat, nemăsurat, nemijlocit, nenumărat, nepăsare, neruşinat, nesaţ, neschimbat, nestins, numărătoare, oacheş, ochire, olan, picioruş, pieptar, pieritor, plîngător, plouat, portiţă, porumbiţă, purceluş, purtător, putinţă, putrejune, rarişte, răci, răcoreală, răcori, răcorit, rămurea, răpejune, rări, rărit, rărunchi, răscula, răsfirat, răspuns, răsuflet, răutate, rechema, reînvia, relua, renaşte, resăltînd, ridicat, rîuşor, românesc, roşi, rotunzi, rotunzit, sălbătăcit, săltăreţ, sărut, scurteică, scuti, secăciune, sîngerat, sori, spînzurătoare, stingere, străvechi, strigăt, stufos, suferitor, suitor, supune, supus, surîzînd, suspinînd, şedinţă, toacă, uitit, umblet, umbros, umflat, urare, urît, urlet, uşurel, uşurinţă, vieţui, vioi, viorică, zbucium . . . . . 2

541—1043. abătut, adaos, adăpost, adăugire, adormire, adormit, adumbri, afînat, aiurit, ajuns, alătura, alăturat, albăstrime, albeaţă, albicios, alegere, alergare, ales, alocurea, aluniş, amarnic, amărî, amărîre, ameţit, amîna, amorţire, anumit, apărător, apăsare, apos, apropiat, apucat, apucătură, apus, arcuş, argintărie, arginti, arinişte, ascultător, asemănător, auriu, batere, bărbătesc, bărbăteşte, bătătură, bisericos, blestemăţie, blestemat, blîndişor, botez, bourel, bulbucat, călăreţ, călărire, călduros, cărăbuş, cărăruşă, căscat, căsnicie, căsuţă, ceruit, cîmpenesc, cîntăreţ, codi,



Scanned with OKEN Scanner

codruţ, comesean, comîndare, contrazicător, costişă, cotitură, crăpătură, creştere, crez, crezămînt, cruciş, crucişa, cruzime, cugetător (subst.), cugetător (adj.), cuibar, cumpătat, cunoscător, cupar, curăţie, curtean, curtenitor, cutezare, cutezător, dărmătură, defăimător, deprindere, descălecător, descoperire, descrescînd, desprimăvăra, desprins, deşărtăciune, deşirat, deşteptat, dezbrăcat, dezlegat, dezmierdător, discîntec, distrăma, dizmînia, domnesc, domnitor, ducere, dumnezeire, dus, facere, făcător, fermeca, fecioresc, ferăstruie, ferecat, fericitor, fermecare, feţişoară, fioros, fireşte, firicel, florărie, floricică, frăţior, frîngere, frînt, frumuşel, fugind, fulgerînd, fumegînd, furculiţă, gălbeneală, gălbini, gălbiu, găuri, gemînd, genuncheat, ghebos, gheţuş, grăunte, grecie, grecotei, greţos, grierel, guraliv, guşat, ierbărie, ieşit, inimioară, inimos, inişte, îmbătare, îmbătrînit, îmblînzit, împăcare, împăiet, împăroşat, împărtăşi, împărţibilitate, împietri, împingere, împlinit, împuţina, înaintit, înălţat, înapoia, înarma, încalea, încalte, încălzit, încărunţi, încetinel, închegat, încheiere, închinare, închinat, închisoare, încorda, încredere, încredinţa, încreţire, încreţitură, încrucişat, încrunta, încurcat, îndărătnicie, îndemnare, îndepărtat, îndrăcit, înduioşat, îndulcit, îndurat, îndurător, înfierbînta, înfiera, înfiorare, înfiorat, înfiorător, înfloritor, înfocare, înfrîna, înfrumuseţa, îngreunat, înjosire, înmormînta, înmulţi, înot, înrădăcina, înrădăcinat, înrîurinţă, însămnătate, înserare, înseta, înspăimînta, înspăimîntat, înşălător, înşelare, înşelăciune, înşeuat, întărire, întîrzia, întraripat, întredeschide, întredeschis, întrerupt, întruni, înţeles, învăliş, învăţător, învineţire, învineţit, învingător, învins, învioşat, înzeit, judecător, juneţe, junghi, lăcrăma, lăcrămos, lăcrimare, lămuri, lărgi, lătrat, lăţi, limpezit, linguşi, linguşire, linişti, lins, luare, luător, lucrător, luminător, lumînărică, lungit, măiestrie, mărginit, mărit, mărmurit, mărunţel, mescioară, minuna,



Scanned with OKEN Scanner

mirat, mişcat, miţos, mîngîios, mînios, mînuşiţă, moliciune, morişcă, muced, mucegai, mucos, mugind, muieratic, murit, muscuţă, mustrare, muşcat, neastîmpărat, nebunatic, nebuneşte, necăutat, necontenit, necuprins, nedescoperit, nedezlegat, nedreptate, neferice, nefericire, nefiresc, neiertat, neîmblînzit, neîmplut, neînsemnat, neînvins, neliniştit, nemişcător, nemîncat, nemîngîiet, neîndurător, nentrerupt, nenturnat, neplăcut, nepreţuită, neprevăzut, nepriceput, nesăţios, nesimţire, nespălat, nestingibil, nestrămutat, nestrăbătut, neşters, neştiut, neudat, neuitat, nimicit, nimicnicie, nodoros, noptos, nourel, numeros, numit, ocheană, orbitor, ostaş, oştean, pacinic, patrulea, pădurean, păgînătate, păgînesc, pămîntean, pămîntesc, părtaş, păs, păscior, păscut, păstoraş, păstoresc, păstrare, pătrat, peire, perit, perzare, petriş, pieptăraş, pieptăna, pieptănat, pierdere, pietricică, pietros, pietrui, pîntecos, plînset, portar, porumbel, preoţime, preschimba, priceput, prins, privighetoare, privitor, prînzi, pureca, purtător, putrezit, rămuros, răpire, răpit, răpitor, răsărire, răsăritor, răsfoi, răsgeme, răsputere, răstimp, răsturnare, răsuflare, răsuflet, răsunător, răsunînd, rătunzire, readuce, recunoştinţă, redeştepta, redispărea, reîncepe, reîntineri, reîntoarce, renăscător, resălta, revedere, românie, rotire, rotitor, ruşinat, sălbătăci, săltînd, sătucean, scăldătoare, scăzut, schimbare, schimbător, scînteiare, scriitor, scris, seamăn, secat, semnătură, seninătate, sfîrticat, sfîşiat, sfîşietor, sîngeros, somnoroasă, somnoroşie, sorbitură, spărtură, spătar, spusă, statornicie, stins, stors, străbun (subst.), străbun (adj.), strănepot, străţese, strecurătoare, strevedea, strigare, strigoi, strîmtoare, strîns, sunător, supărător, supravieţui, supt, surda, şerpui, şezătoare, şirag, ştergar, şters, ştire, tinerel, tomnatec, tors, tort, trecător, tremurat (subst.), tremurat (adj.), trufaş, trufie, trunchiet, tufari, tufărie, tufiş, turburare, turburător, ucigaş, uitător,



Scanned with OKEN Scanner

umbrire, umezi, undoiare, undoiet, undoios, unire, urîciune, urîtor, urmare, urmaş, urmărire, urmărit, urzeală, uscat, uscăciune, uscăţiv, uşurel, văduvie, văduvit, vechitură, veghere, veninos, verişoară, verzui, visătorie, vîlcioară, vîndută, vîntos, vînzător, vrere, vultureşte, zburare, zburător, zecelea, zeiţă, zgîriet, ziuliţă . . . . . 1

După cum se poate observa, mulţimea derivatelor este aproape la fel de numeroasă ca aceea a cuvintelor-bază (1 043 faţă de 1 104). Contrastul cu situaţia din dicţionare, unde, de regulă, la un termen moştenit revin în jur de patru unităţi formate pe teren românesc prin prefixare, sufixare sau derivare regresivă, e din nou izbitor. Uzajul favorizează vizibil cuvintele-matcă, preponderenţa lor aproape insesizabilă, dacă ne referim exclusiv la numărul de cuvinte-titlu, devenind zdrobitoare atunci cînd se iau în considerare frecvenţele absolute de apariţie a acestora. Comparativ cu cele 52 124 de ocurenţe ale cuvintelor-bază provenite din latineşte, derivatele de la termeni de aceeaşi origine nu înregistrează decît 3 477 de ocurenţe adică de 15 ori mai puţine! Aceşti descendenţi indirecţi ai etimoanelor latine, care, priviţi prin lentila măritoare a dicţionarului, păreau să constituie o forţă considerabilă, se văd astfel reduşi la proporţii cu totul modeste. În ciuda unei mase în fond apreciabile, ei nu acoperă decît 5,13% din ansamblul textului.

Examinarea celor două liste de cuvinte ne dezvăluie şi cauza acestui decalaj spectaculos. În inventarul general al frecvenţelor de apariţie, cuvintele-bază ocupă cu precădere rangurile mici, în timp ce derivatele se grupează preponderent în partea de jos a clasamentului. Astfel, printre termenii posedînd mai mult de 80 de apariţii fiecare, întîlnim nu mai puţin de 104 cuvinte moştenite, dar nici măcar un singur derivat. Prima unitate derivată, substantivul *durere,* înregistrează 71 de ocurenţe, iar cea de a doua, verbul *visa,* de-abia 48. În schimb, clasa cuvintelor ce nu apar decît o singură dată include 503 derivate faţă de numai 148 de cuvinte-bază. Fenomenul ne aminteşte de raportul de forţe dintre elementul latin moştenit şi cel neologic „francez" evidenţiat de analizele din capitolele precedente. Remarca funcţionează aici ca o invitaţie de a cerceta dacă nu cumva şi în acest caz operează o legitate de tip exponenţial, asemenea celor ce guvernau raporturile dintre diversele componente etimologice în cadrul vocabularului structurat pe criteriile productivităţii, polisemiei sau numărului de expresii generate. Avem toate datele pentru a verifica



Scanned with OKEN Scanner

o atare ipoteză. Listele precedente ne procură materialul pentru tabelul 21.4 care nu se mai cere comentat, deoarece e structural asemănător cu nenumărate altele deja prezentate în capitolele anterioare.

Reprezentate în figura 21.5, cele 7 puncte furnizate de tabelul 21.4 se dispun aproape coliniar de-a lungul unei drepte avînd panta de

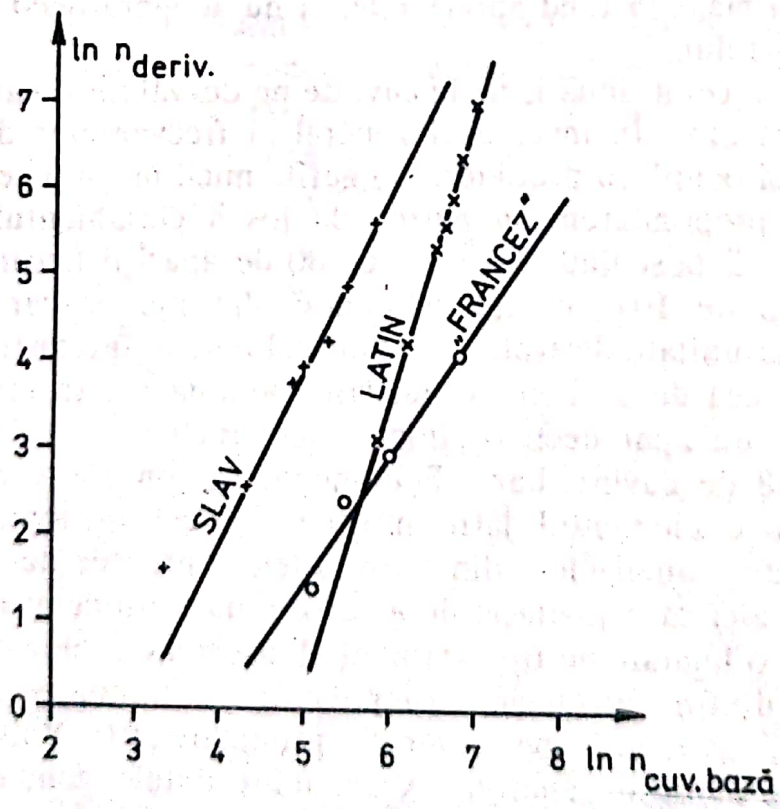
*Tabelul 21.4*

| NR. OCURENȚE | 20 | 10 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| $n_{cuv\text{-}bază}$ | 356 | 503 | 706 | 773 | 856 | 957 | 1104 |
| $ln\ (n_{cuv\text{-}bază})$ | 5,87 | 6,22 | 6,56 | 6,65 | 6,75 | 6,86 | 7,01 |
| $n_{derivate}$ | 22 | 67 | 201 | 256 | 347 | 540 | 1043 |
| $ln\ (n_{derivate})$ | 3,09 | 4,20 | 5,30 | 5,55 | 5,85 | 6,29 | 6,95 |

aproximativ 1 : 3,33. Ne aflăm, aşadar, din nou, în prezenţa unor raporturi lingvistice dominate de o lege exponenţială, mult discutata „liniaritate" (în coordonate logaritimice) confirmîndu-şi încă o dată statutul de principiu universal al organizării vocabularului.

Mai rămîne, desigur, să verificăm valabilitatea celor observate şi pentru alte componente ale lexicului analizat. Nu o vom putea face în toate cazurile, avînd în vedere faptul că cele 5 016 cuvinte folosite de Eminescu reprezintă un fond relativ modest, ceea ce face ca anumite

FIG. 21.5






Scanned with OKEN Scanner

surse etimologice să fie slab reprezentate şi să nu se preteze la o tratare statistică. Constituenţii majori ai vocabularului trebuie însă examinaţi în orice caz. Elementul slav va fi, ca întotdeauna primul dintre ei.

Enumerăm mai jos cuvintele de obîrşie slavă veche din textele eminesciene, cu menţionarea numărului lor de ocurenţe. Mai întîi cuvintele-bază:

| | | |
|---|---|---|
| 1. | iubi | 115 |
| 2. | privi | 111 |
| 3. | val | 80 |
| 4. | vreme | 76 |
| 5. | glas | 67 |
| 6. | veac | 64 |
| 7—8. | izvor, mîndru | 63 |
| 9. | trăi | 50 |
| 10. | trebui | 49 |
| 11. | stîncă | 37 |
| 12. | sfînt (adj.) | 36 |
| 13. | coborî | 35 |
| 14—15. | gît, şopti | 34 |
| 16. | vecinic | 33 |
| 17. | zîmbi | 32 |
| 18—20. | gol, oglindă, zid | 29 |
| 21—24. | drum, noroc, poveste, trezi | 28 |
| 25—26. | babă, icoană | 27 |
| 27—29. | milă, prea, taină | 25 |
| 30. | rînd | 24 |
| 31. | drag | 22 |
| 32—33. | călugăr, vîrf | 19 |
| 34—35. | slab, zadar | 18 |
| 36—40. | bogat, deal, dungă, risipi, scump | 17 |
| 41—43. | iute, ivi, poală | 16 |
| 44—45. | jale, rai | 15 |
| 46—49. | citi, porni, uliţă, zăpadă | 14 |
| 50—54. | dumbravă, jeratic, opri, pod, zidi | 13 |
| 55—60. | aievea, crai, izbi, obraz, pleată, vesel | 12 |
| 61—68. | biet, candelă, clopot, mănăstire, omorî, roi, vrajă, zare | 11 |
| 69—77. | hîrtie, mreajă, peşteră, sclipi, sfînt (subst.) | 10 |
| 78—84. | ceas, crin, goni, învîrti, lipi, stîlp, trup | 9 |



Scanned with OKEN Scanner

85—96. chilie, corabie, crîng, drag, lovi, painjen, război, sfat, tocmai, vină, voie, zori . . . . . 8
97—100. evreu, pleşuv, prund, rasă . . . . . 7
101—105. drag (subst.), lebădă, rob, topi, veste . . . . . 6
106—129. bulgăre, ceată, dragoste, duh, lopată, luncă, mantie, meni, molcom, muncă, năzdrăvan, nevoie, părăsi, plesni, polei, popă, prieten, primi, rac, răspîndi, sticlă, trai, văzduh, vrăji . . . . . 5
130—148. boală, boier, cîrjă, cumpănă, grajdi, grindă, leneş, oglindi, pară, praf, prost, sfîrşi, silă, snop, suliţă, trăsni, var, vîrstă, voios . . . . . 4
149—195. bob, cămin, cîrmă, clăti, clipi, coajă, colibă, colindă, comoară, daltă, eres, gleznă, grămadă, hrăni, iad, idol, isteţ, încovoia, lene, munci, nerod, obicei, obosi, pinten, pipăi, plăti, podoabă, potop, prăpastie, pribeag, răcni, ropot, sădi, schit, sfadă, sfeşnic, sol, strajă, suci, tablă, tămîie, ţapăn, vinovat, za, zdrobi, zvon . . . . . 3
196—245. april, brazdă, buche, ciudă, clăti, coasă, dar, dîră, dovedi, glumă, groază, gunoi, îndrăzni, îngrădi, întîmpina, jeli, lacom, logodi, metanie, mirean, mireasmă, movilă, năpusti, nevastă, norod, organ, pîlc, pîndi, pleavă, plug, poiană, poliţă, prag, proroc, săhastru, stihie, talant, temei, tină, tîrg, tîrî, toiag, treaz, ţigan, urni, vişină, vodă, vrav, vrednic, zimbru . . . . . 2
246—344. arhanghel, arhimandrit, bejanie, ciocan, cîlţi, cîrpi, clocoti, cobe, cocor, cojoc, copită, croi, ctitor, datină, dărui, diavol, dihanie, dojeni, fală, făli, gîngav, gol, har, hrană, huli, iaz, iconostas, irod, iscodi, isprăvi, iute, izvod, împroşca, jertfă, leit, lenevi, letopiseţ, lobodă, mir, mironosiţă, nădejde, năvod, ocol, osîndi, ovăs, pajişte, păzi, perete, pestriţ, pivniţă, poftă, pomeni, postav, posti, potir, pripas, pristol, puhoi, pustnic, puzderie, răzbi, răzvrăti, schiptru, scrînti, scund, sfădi, sfert, sfetnic, slavă, slăvi, sluji, smoală, smuci, sprijin, stareţ, stih, stîrv, stană, strivi, strună, tabără, temelie, temniţă, tîlcui, tîmp, tîrîi, ţintă, vecernie, vifor, vintrea, viteaz, voievod, vornic, vrajbă, vrăjmaş, zăvor, zdravăn, zloată, zmeu . . . . . 1



Scanned with OKEN Scanner

Tabloul derivatelor de la cuvinte provenite din slava veche se prezintă după cum urmează:

| | | |
|---|---|---|
| 1. | privire | 39 |
| 2. | iubire | 34 |
| 3. | clipă | 33 |
| 4. | iubit (subst.) | 23 |
| 5. | tainic | 21 |
| 6. | zîmbet | 19 |
| 7. | miros | 18 |
| 8. | bălai | 15 |
| 9. | atîrna | 14 |
| 10. | sfîrşit | 13 |
| 11. | vecinicie | 12 |
| 12—13. | falnic, povesti | 11 |
| 14—15. | împleti, mirositor | 9 |
| 16—20. | izvorî, mîndrie, mlădios, ostenit, voinţă | 8 |
| 21—25. | iubit (adj.), îndrăgit, jalnic, slăbit, troieni | 7 |
| 26—33. | crăiasă, îndrăgi, painjiniş, podea, sclipitor, smintit, tăinuit, umilit | 6 |
| 34—42. | îngrozitor, întrupat, nenorocire, nesfîrşit, răscoli, smerit, vecie, veselie, zîmbitor | 5 |
| 43—51. | descleşta, despletit, grămădi, nenorocit, pază, privelişte, stropit, tînguios, ulicioară | 4 |
| 52—68. | asfinţit, călugări, despleti, drăguţ, greşală, împrăştia, împrăştiat, încleştat, îndrăzneţ, privinţă, prostie, rumeni, sfătui, spoială, sprijini, stecli, zadarnic | 3 |
| 69—124. | atîrnat, brăzda, călugăriţă, chiliuţă, cititor, clipită, clopoţel, desăvîrşire, dovadă, drăgălaş, goliciune, ispită, iubitor, împodobi, împotrivi, încîlci, încîlcit, încleşta, îngrămădit, înmlădia, întrupa, înţepeni, învoi, jidovesc, lipit, mîndră, mlădia, mlădiet, nepomenit, nerozie, nevinovat, nevrednic, painjini, pletos, pocăinţă, policioară, potrivi, răsplăti, reciti, rėgăsi, risipit, rumenire, săhăstrie, sfiicios, tainiţă, tăinui, tînguire, tînguitor, topit, trăsnet, vădit, văruit, vrăjit, zădărnicie, zdrobit, zdruncinat | 2 |
| 125—261. | anevoie, asfinţi, aţinta, bogăţie, boierie, boierime, bulgăroi, călugăresc, călugăreşte, cernit, cetire, | |



Scanned with OKEN Scanner

cîrpoci, clădire, clopotniţă, cojiţă, cosit, crăiesc, croitoresc, cumpăni, darnic, desăvîrşit, dezgoli, dezlipi, distăina, dobitocesc, drăgosti, drăguliţă, fălos, gîtui, glăscior, glăsui, glăsuire, goană, goli, greş, iscălit, istovit, iuţi, împistrit, încîlcitură, încolţi, încolţit, îndeosebi, îndrăgitor, îngrozire, întocmai, întrupare, înveseli, învinuire, învoială, învrednici, jărtfi, lovire, milos, mitocan, mîndruliţă, muncit, neisprăvit, nevinovăţie, nezărit, obrăzar, obraznic, obrăjel, ocoli, oglindire, osîndă, osîndit, otrăvit, oţelit, pedeapsă, pizmătareţ, pînditor, plesnet, pocitură, poftă, ponorî, ponorîre, pornire, poruncitor, post, potrivit, povăţuitor, pribegi, pribegit, prietenos, prigonire, primejdios, prizărit, prorocire, prosti, prostit, prostuţ, răcnet, răscolit, răsucit, reivi, repovesti, robi, robie, roinic, roire, ropoti, sclipind, sfială, sfiiciune, sfinţenie, silit, slăbi, slăbiciune, solie, spori, sticlitor, sticlos, sticluţă, stihuire, strivit, stropi, şopoti, şopotind, şopotit, şopotitor, tîlcuire, tîmpit, tocit, trăitor, trebuinţă, ţinti, vesti, vestit, vitejie, vîlvoare, vîlvoi, vîsli, vornicel, vremui, zdrenţuit, zgîrcit . . . . . 1

Ambele liste sînt, precum se vede, mult mai sărace decît cele corespunzătoare elementului latin. Într-un fel, ne aşteptam la acest rezultat, dar proporţiile decalajului depăşesc considerabil situaţia cunoscută din despuierea dicţionarelor. Se ştie că D. Macrea a numărat în Dicţionarul Enciclopedic „Cartea Românească“ 2 099 de cuvinte-matcă de origine latină, faţă de 1 466 provenite din slava veche. La rîndul nostru, am identificat în DLRM 1 802 termeni-bază latini şi 1 101 slavi. În primul caz raportul era de 1,43, în cel de al doilea de 1,63. În lexicul lui Eminescu preponderenţa fondului latin e însă mult mai mare, raportul respectiv atingînd valoarea de 3,22 (1 104 cuvinte provin din latineşte şi numai 344 din slavă). Încă mai defavorabilă pentru elementul slav este examinarea frecvenţelor de apariţie a cuvintelor. Dacă cei 1 104 termeni latini totalizează 52 124 de ocurenţe, cele 344 de cuvinte slave se mărginesc la 2 674, adică de aproape 20 de ori mai puţine! Primele ocupă 76,89% din ansamblul textului pe cînd ultimele — de-abia 3,94%. Avînd în vedere că, în general, împrumuturile slave sînt mai puţin productive decît cuvintele



Scanned with OKEN Scanner

moştenite din latină, era de prevăzut ca raportul dintre derivate şi cuvintele-bază să fie mai mic în cazul celor dintîi, ceea ce se şi întîmplă: $\delta_{slav} = \frac{261}{344} = 0{,}759$, faţă de $\delta_{lat} = \frac{1\,043}{1\,104} = 0{,}945$. Totalizînd, unităţile de origine latină (cuvinte-bază + derivate) acoperă 82,02% (55 601 cuvinte-text) din întinderea totală a scrierilor eminesciene investigate, pe cînd cele slave doar 5,10% (3 459 de cuvinte-text).

O întrebare ce s-ar putea pune este în ce măsură situaţia constatată e proprie limbii române în general sau constituie o particularitate a scrisului eminescian. Pentru a încerca un răspuns, vom apela din nou la reprezentările grafice care ne-au facilitat deja în atîtea rînduri accesul către înţelegerea unor raporturi cu caracter mai general dintre diferitele componente istorice ale vocabularului. Cititorul e de pe acum familiarizat cu acest mod de a proceda, aşa încît explicaţiile pe care le mai avem de dat sînt minime. Ca întotdeauna, sistematizăm mai întîi sub formă tabelară datele oferite de cele două liste de cuvinte de sorginte slavă. Tabelul 21.5 e analogul tabelului 21.4, cu deosebirea că acesta din urmă se referea la unităţile de origine latină.

*Tabelul 21.5*

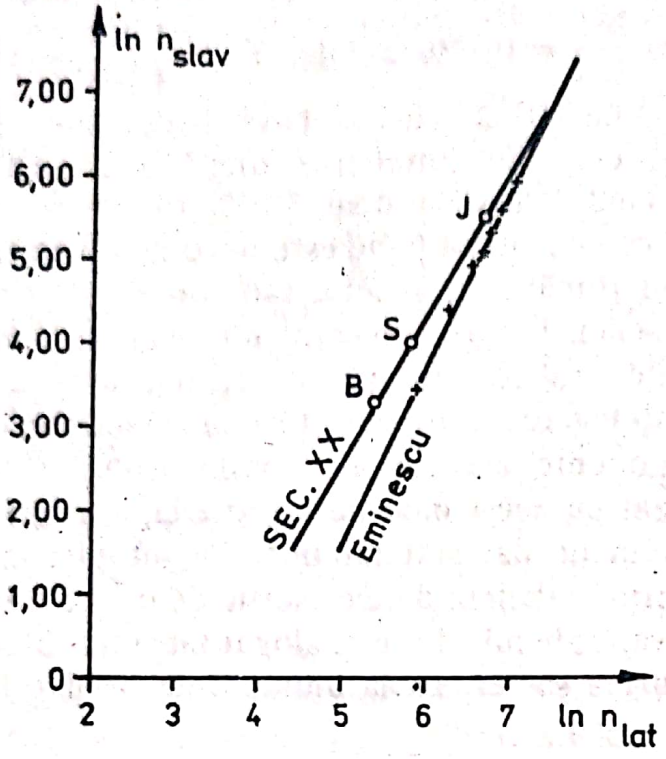
| NR. OCURENŢE | 20 | 10 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| $n_{cuv\text{-}bază}$ | 29 | 77 | 129 | 148 | 195 | 245 | 344 |
| $ln\ (n_{cuv\text{-}bază})$ | 3,37 | 4,34 | 4,86 | 5,00 | 5,27 | 5,50 | 5,84 |
| $n_{derivate}$ | 5 | 13 | 42 | 51 | 68 | 125 | 261 |
| $ln\ (n_{derivate})$ | 1,61 | 2,56 | 3,74 | 3,93 | 4,22 | 4,83 | 5,56 |

Raporturile numerice dintre cuvintele-bază de origine slavă şi latină sînt exprimate, în formă logaritmică, de liniile care ocupă rangul trei în tabelele 21.4 şi, respectiv, 21.5. Cifrele cuprinse în aceste linii reprezintă abscisele şi, respectiv, ordonatele celor 7 puncte marcate cu cruciuliţe în figura 21.6. Tendinţa acestor puncte de a se aşeza pe o dreaptă e manifestă, dovadă că şi în domeniul uzajului legea „liniarităţii" îşi menţine valabilitatea. Constatarea nu ne surprinde, cu atît mai mult cu cît am avut deja ocazia să observăm un fenomen asemănător atunci cînd (v. pag. 263—264) am comparat graficele funcţiei $ln\ n_{slav}$ ($ln\ n_{lat}$) din modelul polisemiei $M_s$ cu datele furnizate de statisticile lui V. Şuteu, C. Maneca, Gh. Bolocan şi A. Juilland. S-a evidenţiat atunci faptul că punctele ce reflectau raportul cantitativ slav-latin erau coliniare, deşi proveneau din excerptarea unor colecţii de texte foarte diferite. Am transpus pe figura 21.6 şi dreapta descrisă



Scanned with OKEN Scanner

FIG. 21.6




de aceste puncte, care poartă aici menţiunea „sec. XX" întrucît toate cele 4 statistici amintite se referă la scrieri din perioada respectivă. Poziţia relativă a dreptelor „Eminescu" şi „sec. XX" e deosebit de grăitoare. Prima se situează deasupra celei de a doua, ceea ce înseamnă că, în mod sistematic, la un număr egal de cuvinte-text de origine latină, scrierile lui Eminescu conţin mai puţine unităţi provenite din slava veche decît textele reprezentative pentru limba literară a secolului nostru selecţionate de cei 4 cercetători mai sus pomeniţi. Semnificativ e şi faptul că în materialul despuiat de C. Maneca şi de A. Juilland figurau şi texte ştiinţifice, în care e de presupus că prezenţa elementului vechi slav e mai discretă decît în beletristică. Concluzia nu poate fi decît una: structura lexicală a operei eminesciene se caracterizează prin privilegierea sistematică a fondului latin în detrimentul componentei slave a vocabularului.

Cît priveşte raportul dintre derivate şi cuvintele-bază, acesta îmbracă şi în domeniul vechi slav forma previzibilă a unei drepte. Am reprezentat-o tot pe graficul din fig. 21.5 pentru a se observa mai lesne că, aşa cum era de altfel de aşteptat, ea se deosebeşte de dreapta corespunzătoare fondului latin printr-o pantă considerabil mai mică. Avînd în vedere experienţa noastră de pînă aici, ne aşteptăm ca înclinarea dreptei ce exprimă relaţia dintre derivate şi non-derivate să

fie cu atît mai redusă cu cît se referă la un material lexical mai „tînăr". Să verificăm această presupunere pe componenta neologică a lexicului eminescian. În acest scop prezentăm în continuare listele cuvintelor de origine „franceză". Termenii nederivaţi sînt, în ordinea frecvenţei lor absolute, următorii:

| | | |
|---|---|---|
| 1. | deveni | 30 |
| 2. | blond | 26 |
| 3. | moment | 23 |
| 4. | surîde | 21 |
| 5. | natură | 20 |
| 6—7. | fin, idee | 18 |
| 8. | dispărea | 17 |
| 9. | portret | 16 |
| 10. | serios | 15 |
| 11—14. | antic, corp, formă, liră | 14 |
| 15. | haos | 13 |
| 16—18. | ruină, stat, univers | 12 |
| 19. | contra | 11 |
| 20—22. | fantastic, tablou, vers | 10 |
| 23—27. | exista, expresie, poezie, roză, voluptos | 9 |
| 28—42. | armonie, cadavru, divin, granit, ideal, mister, notă, obscur, ocean, poet, realitate, sală, statuie, tresări, voluptate | 8 |
| 43—54. | amant, bancă, caracter, damă, existenţă, frază, inunda, melancolic, paloare, patrie, repeta, transparent | 7 |
| 55—71. | admira, alee, contur, diademă, fine, harfă, iluzie, imagine, model, paj, piramidă, reduce, repaos, sens, sferă, sistem, tron | 6 |
| 72—107. | actor, arab, atom, balcon, centru, cochet, crea, crimă, diafan, diamant, enigmă, erou, fantazie, gigantic, glob, inocent, liană, libertate, madamă, mască, melancolie, mistic, mod, nobil, orgie, orologiu, pal, rebel, respira, rezista, sclav, sculpta, seriozitate, simplu, vizionar, zefir | 5 |
| 108—170. | absolut, adora, analogie, animal, azil, bal, bard, boschet, buclă, cadenţă, candid, confunda, contesă, contrasta, cor, cristal, curs, delir, destin, dura, ecou, egal, eternitate, feerie, fibră, fix, forma, formulă, ghirlandă, impresie, individ, infinit, in- | |



Scanned with OKEN Scanner

ocenţă, inspira, intensiv, ironic, încînta, liber, microscopic, milion, misterios, mizerabil, nervos, odă, origine, plan, poetic, presupune, profet, proporţie, ritm, rol, roman, salon, scenă, schelet, secular, simbol, splendid, superstiţios, timid, uman, uragan . . . . . 4

171—245. agresiv, artă, artist, bluză, bravo, bronz, cabinet, cadru, calcul, colonadă, comun, conştiinţă, creion, curtezană, danţ, delicat, demonic, descifra, descrie, dezordine, doliu, element, energie, exemplu, explica, exprima, extraordinar, familiariza, figură, filomelă, fixa, fond, generaţie, grupă, infuzoriu, instinctiv, inteligenţă, intenţie, inventa, întona, luciditate, medita, membru, memoriu, modă, naiv, natal, naţiune, ordinar, paravan, permite, piedestal, piesă, planetă, plebe, preocupa, prezent, profil, proverb, radios, raport, religie, ridicul, sceptic, sceptru, sculptor, sentinţă, silfă, solar, spasmotic, succede, suvenire, teatru, tigru, vegetaţie . . . . . 3

246—418. academie, agonie, agresiune, alarmă, album, amabil, ambră, analog, anima, aplauda, apostat, aranja, ardoare, argumenta, armonios, asirian, aspiraţie, astrologic, astrologie, astru, ateu, atmosferă, aurora, azur, bacantă, basrelief, bizantin, botină, broşură, brun, buduar, bust, capabil, capişon, cita, clasă, coloare, colonie, concentra, condamna, confia, consecvent, consista, constelaţie, conte, conţine, conveni, coordona, corect, creatură, critic, cronică, cultiva, curios, curiozitate, curtezan, dactil, dedica, deja, desemna, despera, dialog, dispune, distinde, dormita, ebenin, egiptean, egoism, elev, emanaţie, epocă, epopee, esenţă, familie, fatal, fenomen, fidel, filozof, filozofie, fizionomie, fluid, (subst.), fluid (adj.), furie, galant, galerie, grimasă, himeric, iamb, ideal, ilumina, impune, influinţă, inimic, insectă, instinct, interesa, intim, invita, irezistibil, lustru, magie, mantă, marchiz, martir, matematic, metafizic, mit, mobilă, monoclu, naivitate, nimb, obiect, ocupaţiune, ondola, onoare, orbită, orcan, organ, oval, papă, pasiune, perfid, pică, plantă,



Scanned with OKEN Scanner

polar, pozitiv, pretexta, prinţ, problemă, profeţie, profum, proletar (adj.), proletar (subst.), promenadă, pronunţa, proză, puls, rasă, reflecta, refuza, regal, regizor, religios, renume, reprezenta, resignat, resignaţie, revoluţie, romantic, sculptură, secret, seduce, servi, sever, sinistru, solitar, sombru, spasm, spectacol, spectator, stindard, sublim, toaletă, tom, torent, torţie, triunghi, troheu, tumultuos, urnă, volan, volum, zvelt . . 2

419—899. abuza, accentua, acomoda, acompania, acorda, acţiune, admite, adonic, advocat, agita, agoniza, alfabetic, alienat, amalgam, amicabil, analiză, aniversară, aparenţă, aparţine, apoplexie, apoteoză, apropria, argument, arhitectonic, arhivă, arid, aristocraţie, aromat, aromă, ascetic, astronom, astronomie, ateist, atelier, aticism, automat, avantagiu, avar (adj.), avar (subst.), aversiune, avizat, banal, barbar, baricadă, batistă, baza, bază, beduin (subst.), beduin (adj.), bibliotecă, biliard, biografie, biped, bizar, blondin, bonetă, briliant, brit, brunet, bucla, butona, caiet, canal, canalie, canar, candelabru, candoare, capitală, caracteristic, carieră, carnaval, cascadă, cauzalitate, cedru, cestiune, chintesenţă, cinism, cipru, clasic, cloroforma, cochet, cochetărie, coincide, coloniza, colora, colorit, colosal, companie, compliment, comprima, concede, concert, conclusiv, confuz, confuzie, congres, consacra, consecutiv, consecutivitate, consideraţie, consola, conspira, constitui, constitutiv, construcţie, consuma, contă, continental, contrast, contribui, controversa, convenţie, convulsiv, copia, copie, copist, coral, corespunde, coridor, corupţie, cran, criptă, cuadrat, culinar, culisă, culpabil, cultură, cumulard, cuplet, curagiu, curent, defensiune, deghiza, delăsa, delicateţă, desemn, deszice, determina, dezgust, dezgusta, diafaniza, diamantic, diluviu, dimensiune, diplomat, disc, discipul, disimula, dispariţiune, disponibil, dispoziţiune, doctrină, domenic, domestic, domina, domino, donjuan, dota, drapa, ducesă, duel, dur, echipaj,



Scanned with OKEN Scanner

echivoc, efect, efemer, efeminat, egalitate, egoist, elastic, elegant, elementar, elibera, elipsă, emoţiune, energic, entuziasma, entuziast, eroic, estetic, excelenţă, excepţie, executa, exemplar, experienţă, expres, extaz, exterior, ezita, faeton, fanfară, favor, fază, feeric, femelă, fenomenal, figurant, filozofic, fineţă, finit, fiolă, flegmatec, floristic, flotă, formal, freme, frenetic, frigian, friza, fulguraţiune, fundamental, galanterie, gală, gamă, gazel, general, generos, genial, gentil, gentileţă, geocentrist, geografie, geometric, glosă, gotic, gratis, guverna, hienă, hipotenuză, idil, idră, ieremiadă, ilot, imagina, imaginar, imaginaţie, imberb, imobil, implora, imposibil, incult, indecis, indefinibil, independenţă, indiferent, individualitate, industrie, infernal, infinitate, infinitezimal, ingenios, iniţială, insomnie, inspiraţie, instituţiune, instrument, interesant, intern, interpret, intrigant, introduce, invers, invoca, ipocrizie, iregular, ironie, înamora, juridic, jurnal, labirint, laic, legendă, leopard, liberal, limbagiu, linţoliu, lugubru, lupanar, lux, luxos, maculatură, magic, maliţie, masă, maşină, mat, matematică, material, matinal, melodios, metafizic, metafizică, metal, metalic, metamorfoză, metempsihoză, meteor, metru, migraţiune, mină, ministru, minut, mitic, mizantropic, monarc, monogram, monoton, monstru, musculos, narcisă, naţional, navă, necromanţie, nefast, nobleţă, nocturn, nonsens, nord, nordic, nostalgie, obscen, observa, observaţie, ocazie, ocupa, orchestru, organizaţiune, orient, ornament, paletă, palm, palton, panoramă, pantalon, paradă, paradis, paragraf, paria, parlament, parlamenta, pasagiu, patruped, pavagiu, pedant, perfecţie, perora, persoană, perspectivă, petrifica, pilastru, placă, plafond, plana, plănta, pleduarie, pleiadă, poem, pomadă, popular, poseda, posteritate, presiune, pretext, pretinde, prezice, prezinta, primitiv, probabilitate, problematic, procesie, profanaţie, profesor, progenitură, proiecta, proprietate, protecţie,



Scanned with OKEN Scanner

prototip, public, pudic, rang, rapsod, raţiune, realiza, reciproc, recomanda, reflecţiune, regiune, regularitate, relativ, reminiscenţă, republican, respectuos, retoric, retrospectiv, reverie, revoca, rezonament, rezonanţă, ritmic, rival, romanţios, roz, saltimbanc, salutar, savant, scald, scandal, schemă, semnificativ, sensibil, senzual, sfinx, sibilă, silf, simfonie, simpatie, simula, simultan, sincer, sirenă, situaţie, slav, societate, sofism, soldat, somnolent, sonet, sonor, spahiu, spaniol, spaţios, statură, strict, suav, subiectiv, subjuga, substanţă, subtilitate, suc, superb, superfluenţă, suprem, surprinde, ştiinţific, tandru, tavernă, temă, tenebră, teribil, teroare, text, timbru, ton, tragic, transforma, transparenţă, transpariţiune, tricolor, tromba, tutui, umoristic, uniform, universitate, uzură, valoare, vals, vast, vegetal, vehemenţă, veninos, venos, vibra, victimă, vil, violenţă, vis-à-vis, vizita, vizită, viziune, voal, vulgar . . . . . 1

În comparaţie cu această listă impresionantă, grupul derivatelor de la rădăcini „franceze" face o figură destul de modestă. Iată-l:

1—4. exilat, încifra, lustruit, parfumat . . . . . 4
5—11. cadenţat, încîntat, preocupat, reflectare, reproducere, suspendat, vibrînd . . . . . 3
12—18. divinizat, inspirat, izolat, mascat, nerealizabil, notiţă, pălind . . . . . 2
19—56. adorare, aspirare, combinare, combinat, deconcertare, dezorganizat, diminuare, diplomaticeşte, dispus, distras, divinizare, emigrare, felicitare, imitînd, încîntare, lustrui, nefinire, neliber, nepavat, nepermis, nepozitivism, neregularitate, nesedus, neviolabilitate, organizat, presupus, pretins, pronunţat, realizare, redeveni, renunţare, respirare, resuvenire, sclavie, scrutător, sinucidere, surprindere, surprins . . . . . 1

Raportul derivate versus cuvinte-bază, aproape unitar în cazul fondului latin, înregistrează aici o valoare extrem de scăzută: $\delta_{fr} = \frac{56}{899} = 0{,}06$. Ca număr de unităţi distincte, prezenţa neologismelor



Scanned with OKEN Scanner

ar fi considerabilă, dar majoritatea absolută a acestora (481 din totalul de 899 de termeni-matcă şi 38 dintre cele 56 de derivate) nu cunosc decît cîte o singură apariţie. Contrastul cu elementul latin, în primul rînd, dar şi cu cel slav, e de-a dreptul izbitor. Numărul mediu de ocurenţe ale unei unităţi nederivate este de 47,13 pentru cuvintele moştenite din latină, de 7,77 pentru împrumuturile slave şi de numai 2,45 pentru neologismele „franceze". Aceeaşi ordine se menţine şi în domeniul derivatelor, însă, de data aceasta, valorile sînt mult mai scăzute: 3,33 pentru fondul latin şi, respectiv, 3,01 şi 1,59 pentru elementele vechi slav şi „francez". Merită notat faptul că, în pofida poziţiei lor modeste, derivatele de la rădăcini latine şi slave posedă totuşi în medie mai multe apariţii decît cuvintele-bază de origine „franceză". Cu numărul lor total de 2 205 ocurenţe, acestea nu ocupă decît 3,25% din ansamblul textelor inventariate. Cît priveşte cele 56 de derivate, ele nu numai că nu totalizează decît 89 de apariţii, dar se prezintă slab şi din punct de vedere calitativ, în sensul sărăciei mijloacelor de formare a lor, majoritatea fiind infinitive lungi şi participii şi doar cîteva derivate autentice cu prefixe *(neregularitate)* sau sufixe *(scrutător)*.

*Tabelul 21.6*

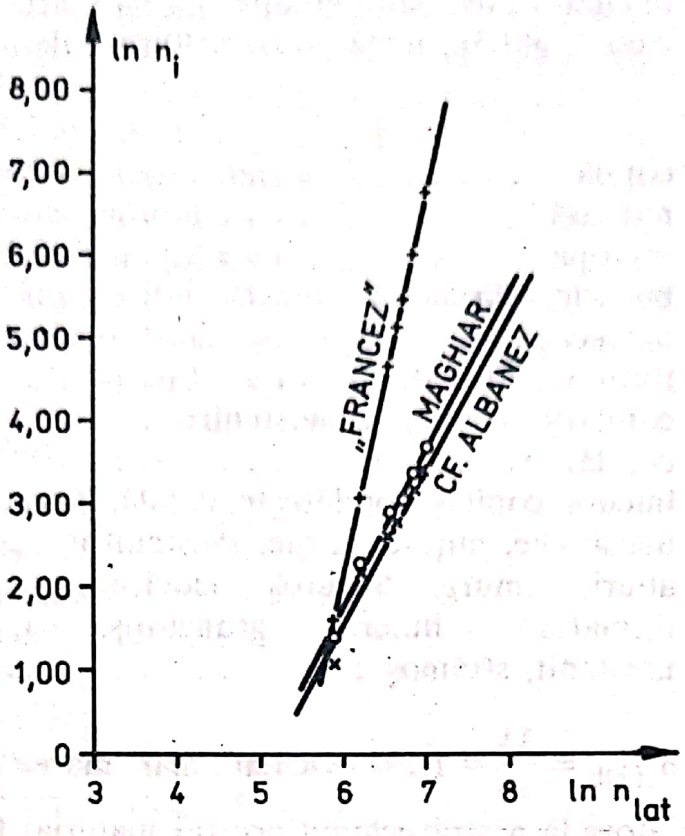
| NR. OCURENŢE | 20 | 10 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| $n_{cuv\text{-}bază}$ | 5 | 22 | 107 | 170 | 245 | 418 | 899 |
| $ln\ (n_{cuv\text{-}bază})$ | 1,61 | 3,09 | 4,67 | 5,14 | 5,50 | 6,04 | 6,80 |
| $n_{derivate}$ | — | — | — | 4 | 11 | 18 | 56 |
| $ln\ (n_{derivate})$ | — | — | — | 1,39 | 2,40 | 2,89 | 4,03 |

Reprezentarea grafică, după tipicul cunoscut, a valorilor logaritmate din tabelele 21.4 şi 21.6 confirmă, pe de o parte, valabilitatea principiului „liniarităţii" (de data aceasta respectat foarte strict, cum e şi normal să se întîmple cu o relaţie de tip probabilist atunci cînd eşantionul considerat e mai bogat; elementele latin şi „francez" prezintă numărul cel mai mare de unităţi distincte şi, ca atare, raporturile lor sînt mai precis descrise decît altele de dreapta din figura 21.7) şi, pe de altă parte, presupunerea noastră că dreapta ce exprimă relaţia cantitativă dintre derivate şi cuvintele-bază în domeniul „francez" ar trebui să posede o pantă inferioară celei ce caracterizează aporturile latin şi vechi slav. Numărul mic de derivate întîlnite face ca în acest caz liniaritatea să nu fie respectată decît de o manieră aproximativă (fig. 21.5).



Scanned with OKEN Scanner

FIG. 21.7




Pentru o bună parte dintre celelalte aporturi, numărul de unităţi disponibile este chiar atît de redus încît trasarea dreptelor caracteristice devine de-a dreptul imposibilă. De aceea ne vedem nevoiţi să renunţăm la examinarea acestora şi ne mărginim la numai încă două componente, întrucîtva mai bogate, ale lexicului eminescian, pe care le trecem telegrafic în revistă în încheierea acestui capitol:

1) Unităţile „comune cu albaneza". Cuvinte-bază:

| | | |
|---|---|---|
| 1. | de | 168 |
| 2. | copil | 59 |
| 3. | buză | 47 |
| 4—6. | brad, jumătate, mal | 15 |
| 7. | fluture | 11 |
| 8—9. | copac, văpaie | 10 |
| 10—11. | mire, pălărie | 8 |
| 12. | vatră | 6 |
| 13—14. | căciulă, gard | 5 |
| 15—17. | brîu, curma, fluier | 4 |
| 18—20. | cioc, măgar, mîţă | 3 |



Scanned with OKEN Scanner

21—24. bucura, ceafă, gata, groapă . . . . . 2
25—30. cioară, ghimp, mînz, moş, strungă, ţap . . . . . 1

Derivate:

1. copilă . . . . . 46
2. mireasă . . . . . 16
3. îngropa . . . . . 11
4—5. bucurie, sfărma . . . . . 9
6. moşneag . . . . . 7
7—8. flutura, sfărîmat . . . . . 6
9—11. copilărie, copilăros, moştenire . . . . . 4
12—13. copilăresc, sugruma . . . . . 3
14—23. bucura, copilaş, copilăreşte, copilă, fărma, gropiţă, măzăriche, mîţesc, moşie, moştenitor . . . . . 2
24—34. aburi, amurg, bucuros, ciocîrlie, copilandru, fărmătură, fluiera, grunzuros, învăpăiere, moştenit, strămoş . . . . . 1

Raportul $\delta_{cf.alb} = \frac{34}{30} = 1{,}13$ e chiar mai mare decît în cazul fondului latin, dovadă a străvechimii acestui material lexical.

*Tabelul 21.7*

| NR. OCURENŢE | 20 | 10 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| $n_{cuv\text{-}bază}$ | 3 | 9 | 14 | 17 | 20 | 24 | 30 |
| $ln\ (n_{cuv\text{-}bază})$ | 1,10 | 2,20 | 2,64 | 2,83 | 3,00 | 3,18 | 3,40 |
| $n_{derivate}$ | 1 | 3 | 8 | 11 | 13 | 23 | 34 |
| $ln\ (n_{derivate})$ | 0 | 1,10 | 2,08 | 2,40 | 2,56 | 3,14 | 3,53 |

2) Elementul maghiar. Cuvinte-bază:

1. gînd . . . . . 68
2. chip . . . . . 46
3. fel . . . . . 29
4. mereu . . . . . 28
5. oraş . . . . . 18
6. neam . . . . . 17
7. gingaş . . . . . 12
8. seamă . . . . . 11



Scanned with OKEN Scanner

| | | |
|---|---|---|
| 9—10. | sicriu, viclean | 10 |
| 11. | chin | 8 |
| 12. | uriaş | 7 |
| 13—14. | întîlni, locui | 6 |
| 15—17. | hotar, meşter, mistui | 5 |
| 18—19. | giulgi, mîntui | 4 |
| 20—22. | bir, guler, ţintirim | 3 |
| 23—29. | bîntui, cotlon, făgădui, locaş, oloi, panglică, pildă | 2 |
| 30—40. | ademeni, bănat, gazdă, ilău, iobag, îngădui, larmă, marfă, şindrilă, uliu, vamă | 1 |

Derivate:

| | | |
|---|---|---|
| 1. | gîndi | 70 |
| 2. | gîndire | 37 |
| 3. | închipui | 12 |
| 4. | închipuire | 7 |
| 5. | mîntuire | 6 |
| 6. | biruinţă | 5 |
| 7—9. | gînditor, mohorît, urieşesc | 4 |
| 10—12. | chinuit, locuinţă, vicleni | 3 |
| 13—16. | chinui, făgăduinţă, locuitor, socoteală | 2 |
| 17—28. | biruitor, chinuitor, dezbumba, gîndit, hotărî, hotărîre, închipuit, înverşunare, înverşunat, mîntuitor, panglicar, vameş | 1 |

Raportul $\delta_{magh} = \frac{28}{40} = 0{,}7$ este ceva mai mic decît în cazul elementului vechi slav, ceea ce concordă cu datele cunoscute privind succesiunea în timp a celor două influenţe.

*Tabelul 21.8*

| NR. OCURENŢE | 20 | 10 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| $n_{cuv\text{-}bază}$ | 4 | 10 | 17 | 19 | 22 | 29 | 40 |
| $ln\ (n_{cuv\text{-}bază})$ | 1,39 | 2,30 | 2,83 | 2,94 | 3,09 | 3,37 | 3,69 |
| $n_{derivate}$ | 2 | 3 | 6 | 9 | 12 | 16 | 28 |
| $ln\ (n_{derivate})$ | 0,69 | 1,10 | 1,79 | 2,20 | 2,48 | 2,77 | 3,33 |

Figurate pe acelaşi grafic (fig. 21.7) cu componenta „franceză", mulţimile de puncte furnizate de tabelele 21.7 şi 21.8 confirmă



Scanned with OKEN Scanner

aşteptările noastre. Dispunerea lor este, şi de această dată, relativ coliniară, iar pantele dreptelor pe care le descriu se dovedesc cu atît mai line cu cît se are în vedere un aport lexical mai vechi.

## 22. În loc de concluzii

Recapitulînd drumul parcurs, în vederea delimitării demersului nostru de alte cercetări complementare sau afine, şi încercînd să caracterizăm cît mai succint conţinutul lucrării de faţă, l-am putea defini sintetic drept „o ipoteză în expansiune". În ultimă analiză, totul s-a redus la dezvoltarea sistematică a unei unice premise, de unde şi impresia, în mare parte justificată, de monotonie a itinerarului, căci exploatarea obstinată a unei singure idei obliga, prin forţa lucrurilor, la sacrificarea multor sugestii colaterale ivite pe traseu, ceea ce a imprimat lucrării un deliberat caracter unidimensional. În plus, experimentul, fiindcă de un experiment e în fond vorba, se putea cu greu sustrage involuntarelor imperfecţiuni inerente oricărei opere de pionierat. În asemenea condiţii, prezumţia formulării de concluzii definitive ca şi ambiţia de a impune drept perspectivă privilegiată un punct de vedere personal, încă nevalidat de confruntarea cu opinia specialiştilor, ar fi fost, desigur, nepotrivite. Căci dacă am cedat tentaţiei de a încredinţa tiparului căutările noastre, nu credinţa în valoarea rezultatelor obţinute ne-a determinat să o facem, ci numai intenţia de a semnala o posibilă cale inedită de abordare a unor teme vechi de cînd lingvistica. Altminteri spus, incursiunea noastră în teritoriul lexicologiei, departe de a se fi vrut o expediţie de cucerire, nu a ambiţionat să fie altceva decît un prilej de verificare a mijloacelor de deplasare, un test funcţional al fiabilităţii vehiculului utilizat. În măsura în care acesta a răspuns corect la comenzi, lucru asupra căruia numai cititorul este în drept să se pronunţe, el se află acum la dispoziţia celor dornici de a purcede la o explorare sistematică a domeniului.

Fireşte, modelarea matematică a structurii etimologice a vocabularului nu putea să nu conducă la evidenţierea unor aspecte încă nesemnalate de cercetarea lingvistică anterioară. Posibilitatea de a opera cu mari mase lexicale şi nu doar cu unităţi individuale, aşa cum ne obişnuise practica lexicologică tradiţională, permite o anumită lărgire de orizont, în sensul facilitării accesului la judecăţi cu caracter



Scanned with OKEN Scanner

mai general. Văzute de sus, lucrurile ne apar, inevitabil, într-o lumină întrucîtva diferită de cea pe care focalizarea pe specimen în dauna speciei o arunca pînă acum asupra fizionomiei lexicului. Relaţiile dintre diferitele aporturi afluente în albia vocabularului românesc au căpătat astfel o înfăţişare, dacă nu cu totul nouă, cel puţin mai complexă şi elastică, vădindu-se că anumite opinii care pînă nu demult treceau drept contradictorii, ireconciliabile, reprezintă observaţii „de nivel" ce, subsumate unei viziuni mai cuprinzătoare, se susţin şi completează reciproc. Acea proprietate neobişnuită pe care am denumit-o convenţional „paradoxul lui A. de Cihac" oferă astfel explicaţia nepotrivirilor aparente dintre rezultatele diferitelor statistici ale vocabularului, permiţînd o mai justă situare a componentelor alogene în raport cu fondul latin moştenit, precum şi o evaluare mai avizată a poziţiei acestuia din urmă în ansamblul elementelor constitutive ale limbii române.

Un avantaj încă neexplorat al noii abordări îl constituie posibilitatea comparării, de pe o bază şi dintr-o perspectivă unitare, a tuturor limbilor romanice, acţiune repetat încercată de lingviştii comparatişti, dar nerăsplătită pînă în prezent decît cu succese parţiale, tocmai din cauză că nu se dispunea de mijlocul de a aduce la un acelaşi numitor universuri lexicale de dimensiuni şi structuri atît de diferite. Modelarea matematică se dovedeşte a fi un asemenea mijloc.

În ceea ce priveşte problema îndelung dezbătută a vocabularului reprezentativ (fond principal, lexic fundamental etc.), trebuie spus că cercetarea noastră, care a pornit tocmai de la analiza structurii acestui compartiment restrîns al limbii, ajunge, în mod paradoxal, să nege legitimitatea conceptului respectiv, care, după opinia autorului acestor rînduri, este nu numai inutil, ci şi susceptibil de a genera destule confuzii. Faptul că unele cuvinte sînt mai importante decît altele constituie un adevăr oricum evident, chiar şi pentru vorbitorul cel mai neinstruit, după cum îndeajuns de limpede e şi faptul că nu se poate trasa o limită netă între un aşa zis nucleu lexical şi masa vocabularului. Statisticile confirmă aceste banale constatări şi cînd o afirmăm nu ne referim numai la inventarele de frecvenţă, ci şi la clasificările pe alte criterii (productivitate lexicală, număr de înţelesuri etc.). Pretutindeni tranziţiile sînt graduale, „elita" lexicală decade treptat din condiţia sa privilegiată, alunecînd imperceptibil în imensa mulţime anonimă a „dezmoşteniţilor" vocabularului. Justificarea operării cu conceptul de lexic fundamental/reprezentativ s-a întemeiat cel mai adesea pe nedemonstrata virtute a acestuia de a oferi imaginea cea mai corectă a raporturilor reciproce dintre sursele istorice ale limbii. Or, tocmai din acest punct de vedere, ideea de fond lexical reprezentativ şi-a



Scanned with OKEN Scanner

dovedit natura falacioasă. Dacă e să reţinem vreo concluzie din lunga şi probabil obositoarea demonstraţie întreprinsă în prezentul volum, ea ar fi aceea că nici o secţiune efectuată prin ansamblul vocabularului nu posedă virtuţi exponenţiale, nu are dreptul de a vorbi în numele întregului. Ca şi în cazul vivisecţiei altor sisteme complexe cu structură asimetrică, anizotropă, lexicotomia furnizează o suită de imagini de o diversitate deconcertantă, pe care numai o viziune globală, de tip holografic, le poate reintegra ansamblului, restituindu-ne faciesul real al organismului supus observaţiei. În fond, către instituirea unei atari perspective au fost îndreptate strădaniiile noastre. Fiind însă vorba de o întreprindere de proporţii ce depăşesc puterile unui singur cercetător, nu s-au putut schiţa aici decît contururile generale ale edificiului plănuit. Aducerea la zi (nu trebuie uitat faptul că, din considerente ce ţin de biografia ideii ce i-a stat la bază, analiza noastră s-a axat pe aspectul limbii române de la jumătatea secolului al XX-lea, reflectînd stări de lucruri deja depăşite de evoluţia vocabularului contemporan), precum şi îmbogăţirea materialului lexical examinat pînă la nivelul de dicţionar tezaur, rămîn sarcini de viitor, a căror rezolvare nu o mai vedem posibilă prin iniţiative individuale, oricît de bine intenţionate şi competente s-ar dovedi acestea. Numai lucrul în echipă, eventual cu profil interdisciplinar, deşi componenta matematică propriu-zisă nu prezintă un grad de complexitate de natură a o face inaccesibilă lingviştilor, oferă şanse de ducere la bun sfîrşit a unui program de anvergura celui sugerat aici. Proiectul reclamă osteneală, răbdare şi chiar abnegaţie, dar socotim că incomparabila creaţie colectivă a românilor care este propria lor limbă merită cu prisosinţă un asemenea efort.

De altfel, reamintirea vechii şi atît de semnificativei omonimii dintre limbă şi popor („Un sultan dintre aceia ce domnesc peste vreo limbă"), ivite din intuiţia ancestrală a solidarităţii esenţiale dintre etnic şi lingvistic, ne întăreşte în credinţa că orice contribuţie cît de mică la cunoaşterea *personalităţii limbii române* poate ajuta la mai dreapta situare a culturii şi civilizaţiei naţionale în contextul european şi universal.

*

* *



Scanned with OKEN Scanner

# Lista referinţelor bibliografice


1. Arvinte, Vasile *Criterii de determinare a împrumuturilor săsesti ale limbii române,* Anuarul de Lingvistică şi Istorie Literară, XVI, Iaşi, 1965, pp. 97—103

2. Bolocan, Gheorghe *Unele caracteristici ale stilului publicistic al limbii române literare* Studii şi Cercetări Lingvistice, XII, 1, 1961, pp. 35—73

3. Brătianu, Gheorghe I. *Tradiţia istorică despre întemeierea statelor românesti,* Editura Eminescu, Bucureşti, 1980

4. Brâncuş, Grigore *Originea consoanei* h *din limba română* Studii şi cercetări ligvistice, XII, 4, 1961, pp. 471—475

5. Candrea, Ion Aurel *Dicţionarul Enciclopedic Ilustrat „Cartea Româneaască"* Editura Cartea Românească, Bucureşti, 1931

6. Coteanu, Ion; Sala, Marius *Etimologia si limba română* Editura Academiei RSR, Bucureşti, 1987

7 Densusianu, Ovid *Histoire de la Langue Roumaine* în Opere, vol. II, Editura Minerva, Bucureşti, 1975

8. Duda, Gabriela; Gugui, Aglaia; Wojcicki, Marie Jeanne *Dicţionar de expresii si locuţiuni ale limbii române* Editura Albatros, Bucureşti, 1985

9. Fischer, Iancu *Al. Graur: Fondul principal al limbii române (recenzie)* Studii şi Cercetări Lingvistice, IX, 1, pp. 127—133

10. Gămulescu, Dorin *Contribuţii la studierea elementelor de origine sîrbocroată ale vocabularului daco-român* Pancevo, 1974

11. Goga, Mircea *Lexicul si structura stilului în poezia lui Octavian Goga* Editura Minerva, Bucureşti, 1989

12. Graur, Alexandru *Încercare asupra fondului principal lexical al limbii romîne* Editura Academiei RPR, Bucureşti, 1954

13. Graur, Alexandru *Fondul principal al limbii romîne* Editura Academiei RPR, Bucureşti, 1957

14. Graur, Alexandru *Tendinţele actuale ale limbii romîne* Editura Ştiinţifică, Bucureşti, 1968

15. Ivănescu, Georghe *Istoria Limbii Române* Editura Junimea, Iaşi, 1980

16. Juilland, Alphonse; Edwards, P.M.G.; Juilland, Ileana *Frequency Dictionary of Romanian Words* Mouton et comp., London, The Hague, Paris 1965

17 Macrea, Dimitrie *Circulaţia cuvintelor în limba română* Transilvania, LXXIII, 4, 1942

18. Macrea, Dimitrie *Compoziţia istorică si tendinţele actuale de dezvoltare ale vocabularului românesc* Cercetări de Lingvistică, V, 1—2, Cluj, 1960

19. Maneca, Constant *Consideraţii asupra frecvenţei cuvintelor în limbă română literară contemporană* Studii şi Cercetări Lingvistice, XVII, 6, 1966, pp. 623—633




Scanned with OKEN Scanner

20. Marcus, Solomon; Nicolau, Edmond; Stati, Sorin *Introducere în lingvistica matematică* Editura Ştiinţifică, Bucureşti, 1966
21. Mihăilă, Gheorghe *Dicţionar al limbii române vechi* Editura enciclopedică, Bucureşti, 1974
22. Pamfil, Viorica *Contribuţii la studiul statistic al lexicului românesc. Structura lexicală a Paliei de la Orăstie* Studii de lexicologie, vol. I, Universitatea „Babeş-Bolyai", Cluj, 1965, pp. 157—232
23. Păun, Gheorghe *Paradoxurile clasamentelor* Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1987
24. Popovic, Ivan *Geschichte der serbokroatischen Sprache* Otto Harrasowitz, Wiesbaden, 1960
25. Reinheimer-Rîpeanu, Sanda *Cuvinte de origine franceză?* Studii şi Cercetări Lingvistice, XXXVIII, 2, pp. 110—114
26. Rosetti, Alexandru *Istoria Limbii Române* vol. I—V (ediţia a IV-a), Editura Ştiinţifică, Bucureşti, 1964
27. Russu, I. I. *Limba traco-dacilor* Editura Ştiinţifică, Bucureşti, 1967
28. Seche, Luiza *Lexicul artistic eminescian în lumină statistică* Editura Academiei RSR, Bucureşti, 1974
29. Seche, Mircea *Schiţă de istorie a lexicografiei române* Editura Ştiinţifică, Bucureşti, 1966
30. Sîrbu, Richard *Antonimia lexicală în limba română* Editura Facla, 1977
31. Şerban, Vasile; Evseev, Ivan *Vocabularul românesc contemporan* Editura Facla, 1978
32. Şuteu, Valeriu *Observaţii asupra frecvenţei cuvintelor în operele unor scriitori români* Studii şi Cercetări Lingvistice, X, 3, pp. 419—445
33. Tudose, Claudia *Vocabularul fundamental al limbii române vechi* în Sistemele limbii, Editura Academiei RSR, Bucureşti, 1970
34. Vascenco, V. *Elementele slave răsăritene în limba română (Periodizarea împrumuturilor lexicale)* Studii şi Cercetări Lingvistice, X, 3, 1959, pp. 394-408
35. Vraciu, Ariton *Limba daco-geţilor* Editura Facla, 1980
36. * * * *Dicţionarul elementelor românesti din documentele slavo-române* (redactor responsabil: Gh. Bolocan), Editura Academiei RSR, Bucureşti, 1981
37. * * * *Dicţionarul explicativ al limbii române (DEX)* Editura Academiei RSR, Bucureşti, 1984
38. * * * *Dicţionarul limbii române moderne (DLRM)* Editura Academiei RPR. Bucureşti, 1958
39. * * * *Istoria învăţământului din România* (redactor responsaabil: Şt. Pascu) Editura Didactică şi Pedagogică, Bucureşti, 1983
40. * * * *Istoria limbii române (ILR)* vol. II, (redactor responsabil: Ion Coteanu), Editura Academiei RSR, Bucureşti, 1969
41. * * * *Istoria limbii române* (coordonator: Florica Dimitrescu), Editura Didactică şi Pedagogică, Bucureşti, 1978
42. * * * *Vocabularul reprezentativ al limbilor romanice* (coordonator: Marius Sala) Editura Ştiinţifică şi Enciclopedică, Bucureşti, 1988



Scanned with OKEN Scanner

# CUPRINSUL


1. Vocabularul ca sistem: formularea problemei. . . . . 5
2. Frecvenţa cuvintelor şi problemele determinării ei. . . 7
3. Locul frecvenţei de folosire printre ceilalţi indicatori ai importanţei cuvintelor. . . . . 12
4. Valoarea şi limitele fondului principal lexical în caracterizarea genealogică a vocabularului românesc. . . . . 20
5. Repartiţia după proveniență a cuvintelor din fondul principal lexical. . . . . 25
6. Raporturile cantitative dintre provenienţe. „Ipoteza liniarităţii“. . . . . 28
7. Relativitatea raporturilor numerice dintre constituenţii istorici ai lexicului. „Paradoxul lui A. de Cihac“. . . . . 45
8. Indicatori ai importanţei cuvintelor: productivitatea lexicală. Corelaţia cu vechimea. . . . . 52
9. Clase de egală productivitate. . . . . 69
10. Dependenţa de origine a puterii generative a cuvintelor. 75
11. Implicaţii, virtuţi şi limite ale modelului matematic. . . 145
12. Datarea împrumuturilor. . . . . 150
13. Confirmarea ipotezei liniarităţii în domeniul productivităţii lexicale. Modelul $M_p$. . . . . 162
14. Explicitarea modelului $M_p$. Raporturile numerice dintre cuvintele de diferite provenienţe. . . . . 168
15. Adecvarea modelului $M_p$ la realitatea vocabularului limbii române. . . . . 182
16. Criteriul polisemiei. Distribuţia pe origini. Ecuaţiile modelului $M_s$. . . . . 186
17. Verificări şi corecţii. . . . . 232
18. Rezultatele aplicării modelului $M_s$. Paralela cu $M_p$. . . 239
19. Numărul de expresii şi locuţiuni ca indicator al importanţei cuvintelor. . . . . 252
20. O propunere de clasificare multicriterială a componentelor etimologice ale vocabularului. . . . . 296
21. Uzajul şi problemele sale. . . . . 311
22. În loc de concluzii. . . . . 362

*Lista referinţelor bibliografice* . . . . . 365




Scanned with OKEN Scanner

Tehnoredactor: Florentina Preda
Coli de tipar:: 23
Bun de tipar: 10 XI 1995

Tiparul executat la tipografia
OLIMP PRINTING SERVICES

Scanned with OKEN Scanner

Cartea de față propune o perspectivă inedită asupra vocabularului limbii române. În contrast cu fragmentarismul ce caracterizează cercetarea etimologică de factură tradițională, ea promovează o viziune radical diferită, în care centrul de greutate al preocupărilor lingvistului se mută de la specimen la specie, de la interesul acordat cazului particular la evaluarea tendințelor generale ale evoluției sistemului lexical. Îmbinând rigoarea științifică cu fluența stilului eseistic (autorul, conferențiar la Universitatea din București, este deopotrivă inginer și lingvist, matematician și teoretician al literaturii), lucrarea se adresează nu doar cercurilor restrânse ale cercetătorilor, profesorilor și studenților, ci și tuturor acelora care cred și simt că ar putea rosti odată cu poetul: „Patria mea este limba română!".

ISBN 973-23-0464-2

6860 +
TL 140
Lei 7000

Scanned with OKEN Scanner